5-2023

The Role Of Multiple Object Views In Early Word Learning: A Dynamic Process

Abigail Julian DiMercurio

University of Tennessee, adimercu@vols.utk.edu

Follow this and additional works at: https://trace.tennessee.edu/utk_graddiss

Part of the Child Psychology Commons, and the Developmental Psychology Commons

Recommended Citation

https://trace.tennessee.edu/utk_graddiss/8153

This Dissertation is brought to you for free and open access by the Graduate School at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.
To the Graduate Council:

I am submitting herewith a dissertation written by Abigail Julian DiMercurio entitled “The Role Of Multiple Object Views In Early Word Learning: A Dynamic Process.” I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Psychology.

Daniela Corbetta, Major Professor

We have read this dissertation and recommend its acceptance:

Jessica Hay, Aaron Buss, Devin Casenhiser

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)
The Role of Multiple Object Views in Early Word Learning: A Dynamic Process

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Abigail DiMercurio
May 2023
I would like to take this opportunity to offer my sincere gratitude to everyone who has helped me during graduate school. I want to thank my advisor, Daniela Corbetta, for her guidance and expertise throughout this dissertation and my graduate career. Daniela provided so much support with my graduate work and also in my life over the last seven years. Without her, I would not be where I am today. I am also grateful to my committee members, Jessica Hay, Aaron Buss, and Devin Casenhiser, for their insightful comments and suggestions on this dissertation. I want to thank my lab mates, John Connell, Asante Knowles, and Duangporn Pattanakul, for their help with data collection. They also provided moral support throughout my graduate career. Their friendship has made these last few years enjoyable and meaningful. I would also like to thank Kymberli Hensley, my undergraduate research assistant, for her reliability coding of my data.

On a personal note, I want to thank my friends from the department. We spent many late nights laughing together and commiserating about graduate school. Without them, this would have been a completely different experience. I also need to mention my partner, Kaleb Kinder. We have shared many days together, writing endlessly and bouncing ideas off of each other, but also escaping graduate school and taking time to live. I am forever thankful for his love and unwavering belief in me. Finally, I could not have finished this degree without the support of my parents, Sam and Debbie DiMercurio, and my sister, Nicole DiMercurio. I cannot stress enough how much my family has been a constant source of encouragement and support throughout my seemingly never-ending education. They have celebrated with me through every milestone and also talked through every setback. I am fortunate to have such an incredible support system.
ABSTRACT

Word learning is a complex process that involves multiple interacting components. One of these components is the motor system. During the first few years of development, the onset of motor skills predicts the development of language skills such that earlier onsets of crawling and walking relate to greater vocabulary sizes. It is thought that this relationship occurs due to a developmental cascade where gaining locomotive skills allows for greater environmental exploration, thus, more opportunities to learn new words. One area of interest in this cascade is object manipulation. Moving objects in a way that creates multiple views is related to larger vocabulary sizes later in development. The present study addresses how object manipulation and creating multiple, varied object views impact novel word learning. Infants aged 20-22 months old participated in two between-subjects experiments. Experiment 1 had no infant-directed object manipulation. In one condition, objects moved in a way that created consistent, fixed views. In another condition, objects moved in a way that created multiple, varied views. Experiment 2 had the same two conditions, but infants created the views independently through object manipulation. The evidence for learning in Experiment 1 was weak. There was no effect of condition, and neither condition was above chance level. In Experiment 2, infants learned novel words at a higher rate when they created varied object views. This condition also had a window of analysis that was above chance level. Notably, there was also a positive correlation between how much infants looked at the self-created varied views during learning trials and their rates of word learning. These results have implications for the relative impact of motor skills and object manipulation on word learning.
TABLE OF CONTENTS

Chapter 1: Introduction ........................................................................................................... 1
  Variability in early word learning ....................................................................................... 3
  The relationship between motor skill attainment and language development ............... 6
  The developmental cascade ............................................................................................... 8
  Object exploration ............................................................................................................. 10
  Object exploration and word learning ............................................................................. 13
  A dynamic systems account for word learning ............................................................... 21
  Present Study .................................................................................................................... 27
Chapter 2: Methods—Experiment 1 ...................................................................................... 31
  Participants ......................................................................................................................... 31
  Materials ............................................................................................................................ 31
  Procedure ........................................................................................................................... 33
  Coding ................................................................................................................................ 36
  Analyses ............................................................................................................................... 38
Chapter 3: Results—Experiment 1 ......................................................................................... 40
  Looking during the learning trials ..................................................................................... 40
  Familiar Words .................................................................................................................. 40
  Novel Words ..................................................................................................................... 41
  Correlations between accuracy and participant factors .................................................. 44
Chapter 4: Summary—Experiment 1 ..................................................................................... 46
Chapter 5: Method—Experiment 2 ........................................................................................ 49
  Participants ......................................................................................................................... 49
  Materials ............................................................................................................................ 49
  Procedure ........................................................................................................................... 50
  Coding ................................................................................................................................ 52
  Analyses ............................................................................................................................... 55
Chapter 6: Results—Study 2 .................................................................................................. 56
  Looking during the learning trials ..................................................................................... 56
  Manipulation during the learning trials ............................................................................. 56
  Familiar Words .................................................................................................................. 56
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Summary—Experiment 2</td>
<td>62</td>
</tr>
<tr>
<td>8</td>
<td>Results—Experiments 1 and 2</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Looking during the learning trials</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Mean accuracy</td>
<td>65</td>
</tr>
<tr>
<td>9</td>
<td>General Discussion</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Fixed object views</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>Varied object views</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Individual differences and learning</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>The role of varied object views in word learning</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Implications</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Limitations and future directions</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Conclusion</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>Appendix</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>Vita</td>
<td>135</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1. Sociodemographic information for participants. .......................................................... 92

Table 2. Example of Trial Ordering for Learning Trials ............................................................ 93

Table 3. Example of Test Trial Ordering .................................................................................. 94
LIST OF FIGURES

Figure 1. Experimental Set-up ........................................................................................................... 95
Figure 2. Stimuli for Familiar Words ................................................................................................... 96
Figure 3. Stimuli for Novel Words ....................................................................................................... 97
Figure 4. Timeline of Learning Trials ................................................................................................. 98
Figure 5. Timeline of Test Trials ......................................................................................................... 99
Figure 6. Looking During Learning Trials (Experiment 1) ................................................................. 100
Figure 7. Familiar Words Mean Accuracy During Windows of Analysis (Experiment 1) ............ 101
Figure 8. Familiar Words Time Course (Experiment 1) ...................................................................... 102
Figure 9. Familiar Words Onset Contingent Plot (Experiment 1) ...................................................... 103
Figure 10. Novel Words Mean Accuracy During Windows of Analysis (Experiment 1) .......... 104
Figure 11. Novel Words Mean Accuracy for Conditions (Experiment 1) .......................................... 105
Figure 12. Novel Words Mean Accuracy for Condition by Window (Experiment 1) ............... 106
Figure 13. Novel Words Time Course (Experiment 1) ....................................................................... 107
Figure 14. Novel Words Onset Contingent Plot (Experiment 1) ....................................................... 108
Figure 15. Novel Words Variability (Experiment 1) .......................................................................... 109
Figure 16. Familiar Words Correlations (Experiment 1) .................................................................. 110
Figure 17. Fixed Condition Correlations (Experiment 1) .................................................................. 111
Figure 18. Varied Condition Correlations (Experiment 1) ............................................................... 112
Figure 19. Looking Correlations with Mean Accuracy (Experiment 1) .......................................... 113
Figure 20. Proportion Looking During Learning Trials (Experiment 2) ........................................... 114
Figure 21. Proportion Manipulation During Learning Trials (Experiment 2) ................................... 115
Figure 22. Familiar Words Mean Accuracy During Windows of Analysis (Experiment 2) .... 116
Figure 23. Familiar Words Time Course (Experiment 2) ................................................................. 117
Figure 24. Familiar Words Onset Contingent Plot (Experiment 2) ........................................ 118
Figure 25. Novel Words Mean Accuracy During Windows of Analysis (Experiment 2) ............ 119
Figure 26. Novel Words Mean Accuracy for Conditions (Experiment 2) .............................. 120
Figure 27. Novel Words Mean Accuracy for Condition by Window (Experiment 2) .............. 121
Figure 28. Novel Words Time Course (Experiment 2) .............................................................. 122
Figure 29. Novel Words Onset Contingent Plot (Experiment 2) ............................................. 123
Figure 30. Novel Words Variability (Experiment 2) ................................................................. 124
Figure 31. Familiar Words Correlations (Experiment 2) .......................................................... 125
Figure 32. Fixed Condition Correlations (Experiment 2) .......................................................... 126
Figure 33. Varied Condition Correlations (Experiment 2) ....................................................... 127
Figure 34. Looking Correlations with Mean Accuracy (Experiment 2) ................................. 128
Figure 35. Manipulation Correlations with Mean Accuracy (Experiment 2) .......................... 129
Figure 36. Proportion Looking for Experiments (Between-Subjects) ....................................... 130
Figure 37. Mean accuracy for Experiments (Between-Subjects) ............................................. 131
Figure 38. Mean accuracy for Window of Analysis (Between-Subjects) ................................. 132
Figure 39. Mean accuracy Condition (Between-Subjects) ....................................................... 133
Figure 40. Mean accuracy for Condition by Experiment (Between-Subjects) ......................... 134
Chapter 1: Introduction

Language acquisition is a complex process that infants learn with minimal instruction. Fenson and colleagues (1994) outline the progression of language development across early infancy. During their first year, infants, on average, only begin to produce words around their first birthday. However, between 16 to 18 months old, there is a rapid explosion in vocabulary development. By 18 to 20 months, infants will begin developing telegraphic speech, combining words forming rudimentary sentences. The rapid progression of language development during this time creates a unique opportunity to study underlying mechanisms involved in word learning.

The process of word learning involves multiple interacting components. These components include cognitive functions, social interactions with caregivers, perceptual experience, and language input. Infants need to learn to identify and categorize objects, recognize the sounds and patterns of language, and understand the meaning and use of words. Notably, infants learn words in a complex, dynamic environment. They encounter a variety of objects, providing them with many opportunities to learn new words. However, the complex environment can present a challenge for infants because they must be able to connect referents to new objects.

Motor skills and physical environmental exploration also play a critical role in language acquisition. Infants can learn about objects' physical properties through object manipulation, which could in turn aid infants in mapping labels onto objects. Understanding the role of the motor system and physical environmental exploration in word learning is crucial for comprehending the complex nature of word learning. By engaging with their environment and
manipulating objects, infants gain a deeper understanding of the physical world and develop essential cognitive and perceptual skills that facilitate language development.

Infants' early gross motor skills tend to predict later language development, such that earlier onsets of sitting and walking tend to relate to greater vocabulary sizes (Libertus & Violi, 2016; Walle & Campos, 2014). However, it is not yet clear whether or how the engagement of the motor system through environmental exploration is related word learning during moments of object exploration. A few studies suggest that object manipulation supports in the moment word learning (DiMercurio, 2019; Pereira et al., 2014; Yu & Smith, 2012). Importantly, these studies suggest that to learn new words effectively, infants need to hear them while the objects are being manipulated. While these studies demonstrate the importance of object manipulation during moments of labeling, it is not yet understood whether or how object manipulation is related to changes in learning.

Some support for object manipulation relating to language development is linked to infants creating varied views of the object. Slone and colleagues (2019) found that infants, who created more varied object views during a free play session, had larger vocabulary sizes six months later. However, this does not address the question of whether object manipulation and varied object views impact word learning. It is possible that the link between larger vocabulary size and varied views is due to more active infants having a larger vocabulary, rather than varied views directly influencing word learning. The present study aimed to understand whether varied object views, either self-created through manual exploration or passively observed, impacts novel word learning for infants. The goals were two-fold: One experiment addressed whether varied object views impact novel word learning compared to singular object views. A second experiment assessed whether the rate of novel word learning depended on infants creating the
varied views themselves through object manipulation or by passively observing the varied object views.

**Variability in early word learning**

Early word learning in the first few years of development is a complex and multifaceted process that involves interacting factors. As a result, there can be considerable variability in development. Horst and Samuelson (2008) identified five tasks infants must do to learn a word, including:

1. Separating the word from the surrounding speech.
2. Identifying the object or concept that the new word refers to.
3. Encoding the word form.
4. Encoding some features or characteristics about the object.
5. Storing the information in a way that allows for later retrieval, with links between the word and features.

Infants pick up this complex process with little to no instruction. In fact, between 1-to-2 years old, infants experience approximately a 300% increase in their productive vocabulary (Fenson et al., 1994). During this period of rapid change, it is important to understand not just individual components of how infants learn words but how they learn words in a dynamic and interactive environment.

During the period of rapid vocabulary acquisition in infants, there is also considerable variability in their productive vocabularies. For example, Fenson et al. (1994) found that at 16 months, infants' comprehension ranged from 75 to 350 words, with an equally broad range in production, ranging from zero to almost 200 words. Furthermore, another study conducted by Frank et al. (2017) revealed that this variability in vocabulary size continues as infants’ approach
two years old, with 22-month-olds still exhibiting significant variation, producing anywhere from zero to over 500 words.

Much of the work on understanding the variability in early word learning has focused on infants' language exposure and the social context of early language development. For example, there is a reliable relationship between socioeconomic status (SES) and language development, such that lower SES infants tend to have worse outcomes in language processing (Fernald et al., 2013) and vocabulary sizes (Hart & Risley, 1995; Hoff, 2003). Hart and Risley (1995) found that children from high SES households were exposed to approximately 30 million more words by age four compared to those from low SES households. This large difference in language exposure was suggested to explain the gap in vocabulary size between children from high and low SES backgrounds. Assuming that the mere quantity of words heard is the sole determinant of language outcomes overlooks the potential impact of the quality of the language environment, the social relationship between infants and their caregivers, and the day-to-day environmental interactions.

A recent perspective on the 30-million-word gap has focused on the complex word-learning process. Hirsh-Pasek and colleagues (2015) investigated the quality of language input that infants receive, specifically during moments of joint attention. Joint attention involves a complex interplay between an infant's interactions with their environment and the shared attention of a social partner, usually a caregiver. The study aimed to investigate language input beyond mere word exposure. Their findings suggest that bouts of joint attention are more closely linked to the quality of language input rather than the number of words heard. While this finding supports the complexity of early vocabulary development, it still focuses largely on the social system of word learning, leaving out other contributing factors.
d’Souza et al. (2017) discuss how early language development is influenced by interconnected developmental mechanisms that were previously thought to be independent. These mechanisms include statistical learning, visual attention, social interaction, and motor ability, and they all contribute to the experience-dependent process of language acquisition. Importantly, d’Souza et al. propose that we, as developmental researchers, must embrace the complexity of early language development. Without doing so, we will not have a sufficient understanding of the developmental processes behind word learning. Similarly, citing the complexity involved in early word learning from multiple interacting systems, Samuelson (2021) recently called for an updated theoretical perspective on how infants learn words, pointing out that many current theoretical models are inadequate.

Word learning is an inherently complex process. Infants face a myriad of challenges while learning new words. One of the most frequently discussed is the problem of poverty of the stimulus. Within the broader concept of poverty of the stimulus, which encompasses all lexical acquisition including grammar, Chomsky (1975) discussed a component related to object-label mapping that pertains to the inherent ambiguity in the world, with many potential referents for any given novel word. While not formally described as poverty of the stimulus, Quine (1964) discusses the issue using a vignette in which an individual is surrounded by people speaking a foreign language. One speaker points to a rabbit and says, "gavagai." The word's meaning is unclear, as it could be descriptive, such as "beautiful," or a command, such as "look." While in the gavagai example, the word's meaning is unclear, it is clear that “gavagai” refers to something related to the rabbit. Another issue is referential ambiguity. In this case, a label is produced, but it is unclear to which object it belongs to (e.g., the rabbit, the grass, the sky, the flowers). While Chomsky (1975) resolved these issues by arguing for innate language-specific mechanisms,
more modern approaches hold that children can overcome these problems with general cognitive mechanisms interacting with the child’s natural environment (Clerkin & Smith, 2022; Pereira et al., 2014; Yu & Smith, 2012). Importantly, they view the infant as an active creator of an ideal learning environment.

In this chapter, I will discuss the obstacles infants encounter when learning to map novel words to novel objects, and how they can overcome them through environmental exploration using the motor system. Additionally, I will introduce a dynamic systems theory framework for understanding word learning. Finally, I will discuss how the present study aims to address the gaps in our current understanding of how the motor system affects real-time word learning.

**The relationship between motor skill attainment and language development**

Several groups of researchers have found a relationship between the attainment of motor skills and subsequent language development. Earlier onsets of motor skills typically predict larger vocabulary sizes (He et al., 2015; Iverson, 2010; Oudgenoeg-Paz, 2014; Walle & Campos, 2014). One of the earliest developmental predictors of this relationship is the onset of sitting upright independently. Libertus and Violi (2016) found that infants who began to sit independently earlier in development had greater receptive vocabulary scores at 10-and-14 months old. A similar relationship is present across the transition from crawling to walking. Walle and Campos (2014) found that in a cohort of infants followed longitudinally there was a subsequent increase in receptive vocabulary scores following the onset of walking. Similarly, in a cross-sectional study, walking infants had greater receptive and productive vocabulary scores than crawling infants of the same age (Walle & Campos, 2014). This same trend over the transition from crawling to walking has been found cross-culturally in a cohort of Chinese
infants (He et al., 2015). Thus, a relationship between the attainment of motor skills and subsequent language development seems to be present.

The previously discussed studies examine the role of gross motor skills and their impact on language development. A similar trend exists for the relationship between fine motor skills and language outcomes. Using a questionnaire on fine motor skills, Alcock and Krawczyk (2010) found a positive correlation between fine motor skills and vocabulary scores in 18-month-old infants. Likewise, the growth of fine motor skills between 6-to-24-months-old predicted expressive language skills at 36 months old, even while controlling for socioeconomic status and general, non-verbal cognition scores (Choi et al., 2018). The trend is even seen in preschool-aged children, where success on tasks designed to measure fine motor skills, such as bead threading and block turning, predicted vocabulary scores (Suggate & Stoeger, 2014). Iverson and Braddock (2011) found a similar trend between fine motor and language skills, as measured by unique words used in a task and utterances per minute. However, the relationship was no longer significant when gesture skills were included as a covariate. Although a relationship between the development of fine motor skills and language outcomes exists, the common mediator may be related to the development of gestures.

Overall, the trend of gross motor development relating to language skills lasts throughout the first few years of development. Wang and colleagues (2014) found that infants who demonstrated advanced motor skills at 1.5 years old also exhibited greater language skills at three years old. Similarly, the onset of walking predicted spatial language at 36 months old (Oudgenoeg-Paz et al., 2015). However, when exploration was added to the model, the relationship between walking onset and language skills fell away, highlighting the role of environmental exploration as a mediator between motor and language development.
The developmental cascade

The relationship between motor skills and vocabulary scores is likened to a developmental cascade where attaining a motor skill unlocks a series of changes in the infant, such as a changing visual field, gaining greater opportunities to explore the environment, and a changing social environment (Bertenthal et al., 1984; Libertus & Violi, 2016). The developmental cascade of exploratory behavior at five months has even been connected to cognitive skills as far out as 14 years in a large cohort of participants, controlling for social skills, behavioral issues, and maternal education (Bornstein et al., 2013). In addition, this study found that the efficiency of exploration, as measured by the variety of objects explored, predicted language-related outcomes, including letter-word identification and passage comprehension (Bornstein et al., 2013).

To understand this lasting cascade, we must first understand how early motor skills contribute to how infants perceive and act within their environment. Marcinowski and colleagues (2019) found infants who need to prop themselves up using their arms explore objects for a significantly shorter amount of time than infants that can sit independently. Following the change in sitting skills, object exploration fundamentally changes, which could relate to the developmental cascade between sitting independently and receptive vocabulary scores. Greater duration of object exploration after the onset of independent sitting only benefits infants if objects are within their reach. One of the next major motor skills to develop, crawling, fundamentally changes environmental interactions. Thurman and Corbetta (2017) found that as infants began to crawl, they engaged in greater environmental exploration and traveled further distances in the room.
Similarly, the onset of independent walking is associated with many changes among several developmental systems. For example, infants' visual experiences when crawling compared to walking are fundamentally different. Kretch and colleagues (2014) found great differences in the field of view that crawling and walking infants experienced. Crawlers tended to see more of the floor and had to halt their locomotion to sit upright to see their caregivers. In contrast, walkers had an expanded visual field that mostly included their caregivers (Kretch et al., 2014). This finding has implications regarding the developmental cascade of motor and language development. For example, walking infants can more easily connect seen distal objects and labels that may be produced during exploration.

The visual field is not the only area to significantly change with the onset of walking. Walking affords greater opportunities for exploration within the environment, mediating language outcomes (Oudgenoeg-Paz et al., 2015). Karasik et al. (2011) followed infants from 11 to 13 months old and measured motor skills and bouts of exploration in an environment with a mixture of proximal and distal objects. At 11 months old, all infants were crawling and carried objects similarly. Two months later, half of the cohort transitioned to walking. The walking infants traveled farther in the room to distal objects and engaged in more object-carrying behaviors (Karasik et al., 2011). A similar trend occurs when crawling infants are given artificial, upright walking experiences. Gustafson (1984) compared two groups of crawling infants, one in their normal locomotion and the other in a walker, giving them upright locomotion. The infants in the walker traveled to a distal object that was not explored by the crawlers (Gustafson, 1984). With the onset of walking, infants travel further in the room and interact with previously unexplored objects.
Changes in locomotion also impact the infant's social experiences while navigating their environment. Walle and Campos (2014) found that the overall language environment, concerning the number of words and referents, during a free play session did not differ between crawlers and walkers. However, other studies show that the quality of social experiences differs between crawlers and walkers. For example, Karasik and colleagues (2011) showed that walking infants traveled to distal objects more than crawling infants. They also increased their object-carrying behavior for these distal objects due to having both hands free during upright locomotion. Subsequently, the increase in object carrying resulted in increased bouts of social bids to the caregiver (Karasik et al., 2011). The specific type of social bid that increased with walkers were moving bids, where the infant would go to a distal object and carry it back to their caregiver. These instances of moving bids result in changes in the caregiver's response. Stationary bids (i.e., when an infant presents an object to their caregiver that was not retrieved from a distal location) often elicited affirmations from the caregiver. In contrast, moving bids elicited more action directives (Karasik et al., 2014). Action directives could relate to a wider range of vocabulary used than affirmations, which could support the differences seen in vocabulary scores between crawling and walking infants. Thus, the overall quantity of words may not change between crawlers and walkers. Still, the quality of words, as shaped by environmental exploration, could explain changes in language development as infants transition across motor milestones.

**Object exploration**

The previously discussed literature focused on the changes seen in multiple developmental systems as infants transition through different postures and modes of locomotion and their subsequent impact on language development. One of the most influential areas of
development to change is environmental exploration as well. For example, as compared to infants without independent locomotion, crawlers travel further in a room and engage in more interactive bouts (Thurman & Corbetta, 2017), while walkers engage in more object-carrying behavior, thus eliciting different responses from their caregivers as compared to crawlers (Karasik et al., 2011, 2014).

There are shifts in how infants explore objects dependent on gains in other areas of motor development. For example, Thurman and Corbetta (2019) found that while transitioning to crawling, infants primarily used a sitting posture to engage in fine motor manipulation of objects in a playroom. However, as infants became more mobile, they decreased the amount of fine motor manipulation while sitting, instead using a posture such as squatting or kneeling to engage in fine motor manipulation. Similarly, as infants transitioned from crawling to walking, object interactions became more passive (e.g., holding) while sitting or standing, with squatting or kneeling being the most frequent posture for object manipulation. During this time, infants more frequently changed their posture and switched the object they were interacting with.

The posture between infant-mother dyads also shifts alongside changes in motor development. Schneider and colleagues (2022) found that from 3 to 12 months old, infant-mother dyads shifted from primarily face-to-face sitting postures to postures with better spatial co-orientation (e.g., sitting at a right angle with each other). This shift occurred when infants could sit independently without being propped up by an external object, regardless of age. The shift in how infant-mother dyads position themselves was related to changes in how infants engaged with objects in front of them. Before independent sitting, typically, it was the mother initiating object interaction. After the onset of independent sitting, infants more often initiated object interaction by bringing objects in and out of the interactive space.
Moreover, the specific type of object exploration shifts, as does infant posture. Soska and Adolph (2014) found that when infants were sitting, compared to supine or prone, they were more likely to engage in complex object manipulation. This included rotating the object and passing the object from hand to hand. Infants would also look at the object while manipulating it more often while sitting or prone than supine. Herzberg et al. (2022) examined changes in object exploration across the transition from crawling to walking in 13-month-old. Across this motor milestone shift, the number of long object exploration bouts decreased such that crawlers engaged in more bouts of object exploration that are longer than walkers. This could suggest that walkers are interacting with more objects for shorter bouts. Overall, these findings suggest that infants’ object exploration is influenced by gains in their motor development, leading to changes in posture and object manipulation strategies. Therefore, environmental exploration and perhaps object exploration is central to changes seen in language development across these motor milestones.

Understanding the development of early object exploration is important because there is a direct link between object exploration and communication skills, similar to the findings relating to broad environmental exploration and vocabulary development. Orr (2020) found that exploration directly impacted object play, subsequently impacting gesture and speech production. Orr proposed that object recognition is the central mechanism driving this relationship. Playing with the object involves recalling information related to the object and provides a link with object naming. This supports trends seen in prior work (Oudgenoeg-Paz et al., 2015) that exploration is associated with language development. However, Oudgenoeg-Paz et al. is the first to demonstrate that manual object exploration specifically mediates the association between environmental exploration and language development.
Object exploration and word learning

Although there is ample support for the connection between motor development and language learning through object exploration, very few studies directly link motor skills as related to word learning. In adults, object manipulation during labeling has been shown to benefit word learning when the movement is meaningful (Fuhrman et al., 2021). In this task, adults proceeded through three within-subject conditions in which they were presented with objects labeled in Finnish. The adults had no Finnish proficiency and the task took place in a virtual reality setting. The conditions were: watch-only, where the participant repeated the object’s label while watching a performer move the object; irrelevant movement, where the participant manipulated an object in a way that was irrelevant to the objects typical movement (e.g., tracing a half circle with the object); and a manipulation movement where the participant moved the object in a meaningful way (e.g., putting on a hat). The condition with the highest accuracy was the manipulation movement condition. In this case, object movement helped ground an abstract object with its novel referent.

The benefits of object manipulation in language learning are similar to those proposed by the Physical Interaction as Grounding for Language Transformers (PIGLeT) model (Zellers et al., 2022), which posits that an embodied agent can learn labels by grounding abstract language in the physical world. By giving the agent a joint linguistic model of form and meaning, it can reason about described actions and assign corresponding, correct labels. For instance, when the agent is given the input “pan” and “heatUp,” it learns the labels “hot” and “cooked.” Overall, these findings highlight the importance of physical interaction in language learning, as it allows for the abstraction of language to be grounded in concrete, tangible experiences.
In-the-moment visual properties of objects, as changed by the infant during bouts of object exploration, impact whether or not novel object labels are learned. Pereira and colleagues (2010) found that the visual experience of learned object-label pairs were inherently different compared to unlearned object-label pairs. Learned objects dominated a larger percentage of the infant’s view and were more centered in the visual field. Furthermore, during a play session between infant and caregiver, Yu and Smith (2012) found that infants better learned a novel object label when the caregiver spontaneously provided the label when infants held the novel object. They termed this “embodied attention” because the infant's actions, through object manipulation, created optimal moments of visual attention during the labeling moment. During the labeling moment, the object dominated a majority of the visual field and physically obscured other potential referents on the table. Therefore, the infants' actions in their environment help overcome the issue of referential ambiguity in early word learning.

However, the role of object manipulation may benefit word learning beyond creating moments of embodied attention. In a lab task, DiMercurio (2019) found that infants who did not learn an object during a visual-only object-label presentation learned novel object-label pairs better if they manipulated the object while the label was produced. Additionally, infants who created more object movement and watched the object move during the labeling moment learned the novel object-label pairs at a greater rate. In contrast to Yu and Smith (2012), the objects were presented one at a time, so there were no other potential referents. Therefore, the benefit of actively exploring an object during a labeling moment may extend beyond the role of embodied attention.

When infants engage in object manipulation, they can move toys around in a way that creates multiple and varied views. The differing views that object exploration affords may
explain the link between object exploration and word learning. For example, Slone et al., (2019) brought infants into the lab at 15 months old and observed their object manipulation. Infants who manipulated objects in a way that created a greater number of varied views had greater vocabulary growth over the next six months. Although this is not a causal link between object views and word learning, it demonstrates that infants with a greater propensity to manipulate objects in a way that affords more object views also have greater vocabulary scores. Thus, there may be a link between object views and word learning.

**Object views**

The way infants explore objects, and create differing object views, fundamentally shifts over the first two years of life. Smith (2013) proposed a connected developmental pathway, linking fundamental skills like sitting and object manipulation to the eventual outcome of early noun learning. This pathway heavily emphasizes how visual and manual exploration create moments of stabilized attention, thus aiding in object recognition and word learning. Pereira and colleagues (2010) found that infants around 12 months initially manipulate objects in a way that creates more disorganized, sporadic views of the object. However, two-year-old infants manipulate objects in a more adult-like way, showing a preference for planar object views (views that go along with the perpendicular axis to the front of the object).

In adults, there is evidence that active exploration of objects impacts later visual object recognition. Harman et al. (1999) had participants rotate objects on a computer screen up to 360 degrees for 20 seconds. The rate at which they rotated the object, fast or slow, was not controlled. The control condition was observing another participant's exploration of the object. Therefore, they could control the impact of passively observed object views against actively created object views. The active object view condition resulted in higher rates of later object
recognition. Thus, the active experience, controlled by the individual, led to better object recognition. In particular, the adults tended to create different object views while rotating along the planar axis. While this particular study does not examine learning of object labels, it sets up the important role of actively creating object views in later learning outcomes.

For infants, the development of a preference for planar views coincides with their ability to recognize familiar stimuli given sparse visual information (James et al., 2014). Therefore, it could be that planar views could assist infants in visual object recognition. During this shift of object recognition using sparse visual information, infants typically experience rapid growth in their vocabulary size (Yee et al., 2012). In fact, James and colleagues (2014) found that infants (aged 18 to 24 months old) who had a developed preference for planar views and could recognize shapes given sparse visual information, had higher vocabulary scores as measured by the MCDI compared to their peers who created fewer planar views. Interestingly, these two groups did not differ on any other factors including age, their scores on the general metric of typical development, or even the overall time spent manually exploring the object.

James et al. (2014) proposed that having a bias for planar views when manually exploring objects, allows infants to better integrate multiple views of the object. If the views are non-planar, then the view the infant experiences is more disorganized, and thus more difficult to integrate. The integration of multiple views could then allow the infant to better learn the label for a novel object. Finally, knowing the object label and experiencing multiple views of it, allows for building object categories, which feeds into object recognition given sparse visual information.

Infants, through their object manipulation, create ideal circumstances to view objects in a way that promotes object recognition and novel word categorization. Yee et al. (2012) found that
infants perform better on an object-recognition task before extending novel labels onto objects of similar shapes. The shape bias entails infants categorizing novel nouns by shape and then generalizing the referent onto a novel object of that same shape, but not texture or color (Landau et al., 1988). Generalizing novel labels onto same-shaped objects is an important feature in early word learning because infants will rapidly generalize learned words onto novel instances of a same-shaped object. Therefore, Object manipulation is linked with recognition and inherently linked with a shape bias, which highlights the significance of recognizing and categorizing objects based on shape during early word learning.

**The role of label timing**

Object views may be important to novel word learning, but another factor is involved in the complex word learning process—the timing of when a novel word label is provided. Successful word learning for infants in a natural setting is challenging due to the ambiguity of what could be labeled in a cluttered scene. For instance, in a playroom with many potential toys, a caregiver might label a "ball." Still, the infant may not pay attention to the labeled object, depending on when the label is provided. Thus, the timing of novel label presentation is crucial for successful word learning. Yu and colleagues (2021) investigated how infants respond to referential ambiguity when objects are labeled in a naturalistic setting. The study used typical objects found in a playroom rather than novel shapes, and the goal was to observe where infants looked when parents provided a label (e.g., "it's a car"). The findings revealed a bimodal distribution in infant gaze behavior: they either looked directly at the correct referent or did not look at the labeled objects at all during the labeling. This suggests that infants face significant uncertainty in determining the correct referent when provided with labels, and they need to overcome this challenge for successful word learning. Furthermore, looking initially at the
incorrect referent and eventually directing gaze to the correct referent occurred rarely. Thus, Yu and colleagues (2021) describe an "all or nothing" approach to referential ambiguity in naturalistic scenes. The authors propose that we need to shift how we approach referential ambiguity because infants are not sampling potential correct referents, they either landed on the correct object or did not. Therefore, when the label is provided is crucial for better word learning.

In the previously mentioned studies (DiMercurio, 2019; Yu & Smith, 2012), not only was the manipulation of the object critical to learning the novel word but when the label was produced also impacted if the word was learned or not. DiMercurio (2019) had two within-subject conditions involving manual object exploration that differed in when the novel label was produced. In one condition, the Out-of-Sync condition, the labeling was asynchronous to the object manipulation. In the other condition, In-Sync, the labeling was synchronous with object manipulation. Additionally, there was a visual-only condition where infants saw a novel object and heard the novel label, but never manipulated the object. The results of this study showed that infants who did not learn the label in the baseline, visual-only condition, showed higher rates of learning in the In-Sync condition, but not in the Out-of-Sync condition. Therefore, there was a boost in learning when there was a synchronous moment between holding and looking at the object, while hearing the novel label.

Furthermore, DiMercurio (2019) also accounted for how infants manipulated the object and found that infants who attended to their object movements better learned the novel label only in the In-Sync condition. This same trend was not seen for the Out-of-Sync condition, nor did the amount of manipulation without attending to the object predicted learning. These results suggest
that the moment of intermodal synchrony between holding the object, seeing the object, and hearing the label was crucial in whether or not infants learned the novel object-label pairing.

Similarly, Schroer and Yu (2022) found that timing of label presentation during object manipulation is crucial to learning. In their naturalistic play task, infants were engaged in bouts of multimodal attention, defined as looking at an object while holding it. The label for the object was better learned when it was provided during the bout of multimodal attention or 3 seconds before or after. This finding reinforces the importance of intermodal synchrony in successful word learning. Together, these studies highlight the critical role of object manipulation and intermodal synchrony in learning novel words.

The moment of object labeling that features synchrony between multiple modalities has been highlighted before between caregiver-infant interactions with novel stimuli. Infants, in particular, are sensitive to synchrony across multiple modalities (i.e., a bouncing ball visually hitting a surface, paired with the sound of the ball hitting the surface), and prelexical infants better learn single syllable word-object pairs when the object moved in synchrony with the label being produced (Gogate & Bahrick, 2001; Lewkowicz, 1996). Additionally, Gogate and colleagues (2006) had caregivers play with their infants to teach them the novel label. When caregivers spontaneously produced moments of temporal synchrony, their infant better learned the novel label. An example of this temporal synchrony would be moving the novel object upwards if the inflection of the voice also moved upwards while labeling the object. Infants’ attention toward the object mediated this effect. The temporal synchrony that the caregivers created resulted in greater amounts of attention toward the object. Thus, the labeling moment paired with the object movement created heightened moments of attention, facilitating learning.
While Yu and Smith (2012) highlight the role of embodied attention during the labeling moment, attention and interest in the stimulus are important for in-the-moment word learning. Dunham and colleagues (1993) had infants in two between-subject conditions that differed in when the label was produced. In the first condition, infants played in a room with various novel stimuli. An experimenter would only label the novel object after bouts of infant-directed play with the object. Thus, the label always followed naturalistic bouts of attention toward the object. This condition was used to yoke the number of labels in the other condition, so the mere quantity of labels used was consistent between the two conditions. In this condition, the experimenter would redirect the infants’ attention toward the novel object and then provide the label. Dunham et al. (1993) found that regardless of many times the label was produced, infants better learned the novel object-label pair in the attention-following condition. While the label was not provided during bouts of manual exploration in this study, it still highlights the role of attention toward novel objects in learning labels.

These studies demonstrate that infants are highly sensitive to the moment novel labels are produced, and that visual attention to the object alone is insufficient. Instead, in the moment, word learning is a highly dynamic process involving three factors that impact novel object-label mapping: attention, object exploration, and label timing. These three factors interact, feeding into how well infants learn a novel object label pair. Framing novel word learning within the dynamic systems theory can better explain how infants approach novel situations where they are presented with object-label pairs and how they benefit from multiple systems converging during bouts of labeling.
A dynamic systems account for word learning

Thus far, it has been demonstrated that gains in motor skills during infancy create a developmental cascade that contributes to gains seen in vocabulary sizes. As infants gain motor skills, they engage in more bouts of environmental exploration (Thurman & Corbetta, 2017), seek out caregivers more in social bids involving objects (Karasik et al., 2014), have wider visual fields as they gain upright locomotion (Kretch et al., 2014), and engage in more object exploration (Marcinowski et al., 2019). Therefore, when studying the development of a complex system, such as language, we must consider other areas of infants’ experience. A theory to account for these experiences is the dynamic systems theory.

Studying development from the dynamic systems theory perspective assumes that development is non-linear, complex, and self-organizing (Thelen, 1992). Importantly, dynamic systems theory moves away from a modular view of different systems, like motor and language, and instead recognizes the interwoven connection between multiple domains of development. Importantly, dynamic systems theory views cognition as embodied, such that intelligence or knowledge emerges from our real-time interaction with the physical world using sensory and motor skills (Smith, 2005; Thelen & Smith, 1996). The dynamic systems theory moves the field away from the view that acting and perceiving are outputs of what the brain tells the body to do and instead embeds the brain within the body, within a complex environment (Corbetta, 2009). In this view, acting and perceiving are knowing. Dynamic systems theory holds several basic principles for how it views development, and these can be applied to the tenants of early word learning.
Language as a softly assembled process

One of the first understandings of development through dynamic systems theory is that behavior and cognition are softly assembled. Behavior arises from continued interactions between individuals and their environment (Thelen, 1992). Soft assembly does not view development as hard-wired or innate. If behavior emerges from real-time interactions, the brain is not viewed as the “controller” of the body’s action (Spencer et al., 2011). The understanding that cognition interacts with perception and action is not a novel idea. Piaget (1952) highlighted the importance of perception and action in the development of cognition. Gibson (1988) explained that spontaneous exploratory activity is the framework for the development of cognition. An example of early exploratory behavior is the spontaneous kicking behavior in infants. Infants kick while in the supine position at random rates, and the behavior appears disorganized. In a cleverly designed study, Rovee and Rovee (1969) attached a cord to an infant’s ankle, allowing an overhead mobile to move in response to the leg with the cord kicking. Compared to controls, infants in this condition doubled their baseline kicking rate. In the moment, due to their exploratory behavior, infants learned that their kicking affected the environment and changed their behavior. Furthermore, Angulo-Kinzler and colleagues (2002) followed a similar paradigm. However, the mobile would only move with a specific kick pattern. In the moment, young infants could assemble their spontaneous behavior to be highly organized to get the reward of a moving mobile.

Language development is context-dependent and can also be viewed as softly assembled through in-the-moment exploration within the environment. One view of language acquisition is that it develops through innate, domain-specific mechanisms (Chomsky, 1986). However, recent research suggests that language development is not solely predetermined but instead arises from
real-time exploratory behavior. The relationship between object manipulation and word learning provides an example of this soft assembly. Studies have shown that infants better learn an object-label pair when they manipulate objects and hear novel labels simultaneously (DiMercurio, 2019; Schroer & Yu, 2022; Yu & Smith, 2012). This moment of multimodal attention leads to an ideal learning environment for a novel word label, and it is the infants' own object manipulation that creates this moment. In essence, the way infants manipulate the object gives rise to the label-object pairing, allowing infants to assemble the appropriate views and experiences necessary to learn a word in real-time. Therefore, word learning is softly assembled through exploratory behavior, which provides infants with more opportunities to learn words throughout their environment.

**Stable attractor states of language**

Another important tenant of dynamic systems theory is understanding that behavior, while variable and dependent on context, tends to move towards a stable state, called the attractor state (Thelen, 1992). Attractor states can be viewed as a series of hills and valleys. For example, we can consider a ball resting in a shallow compared to a deep valley. The shallow valley does not require a lot of energy to push the ball out, but a deeper valley does. The deeper the valley, the more stable the behavior is. When there is a shift between attractor states, that is when new behaviors emerge. The A-not-B task provides a great example of attractor states. The canonical version of the task, developed by Piaget (1952), requires infants to view an object hidden in an “A” location multiple times and search for the object each time. Infants tend to perform well on this part. However, after several A trials, the experimenter switches the hiding location, and the infant is cued to search in the second “B” location. Infants tend to perseverate
and search for the hidden object in the A location again. Piaget (1952) attributes this error to weak or incomplete representations of the object out of view.

The dynamic systems theory approach frames the repetitive A reaches for the hidden object as becoming the attractor state for the task. Diedrich and colleagues (2000) recorded the kinematics of infant reaches during the task. They found that if infants demonstrated reaches that became increasingly similar during the A trials, they were more likely to make the error during the first B switch trial. However, if infants’ reaches during the A trials were more dissimilar and disorganized, they were less likely to make the error. Additionally, the number of A trials matters; the greater number of A trials, the more likely infants are to perseverate. The repetitive reaches toward the A location essentially builds, through experience, a deeper, more stable pattern of behaviors (Thelen et al., 2001). Because the nature of the task is relatively difficult for young infants that are still inexperienced with reaching, overcoming the stable reaching state they have built up over the task’s timescale becomes difficult. Thus, they remain in their stable state of reaching toward the A location.

An example of an attractor state in language development is seen in the emergence of the shape bias, meaning that children are more likely to generalize novel nouns to similar shapes than objects with different shapes but the same texture (Landau et al., 1988). This bias begins fairly weak. Younger infants will accept a small degree of shape change when generalizing novel nouns. However, adults do not. Smith and Gasser (2005) proposed a word learning process where infants first learn to map novel labels onto novel objects. Then, they recognize, for example, that cups tend to be cup shaped. After that, there is a second-order generalization that objects they label tend to be shaped a certain way. Because of this reliable trend, infants more reliably attend to the shape property when learning novel labels.
Additionally, Perry and Samuelson (2011) found that the number of shape nouns in a child’s vocabulary predicts the likelihood of a shape bias. Conversely, toddlers with more non-shape nouns in their vocabulary lean towards a texture bias. Using Smith and Gasser’s (2005) framework, the increasing reliability of toddlers’ ability to recognize that labels generally refer to shapes becomes the stable attractor when assigning labels to objects in ambiguous situations. If infants have more variability in their exposure to shape labels, just like more variability in the A trial reaches, they are less likely to have a strong shape bias. This is seen in a study with infants who had larger non-shape noun vocabularies (Perry & Samuelson, 2011) and in a study with an experimental manipulation where infants are trained using the same novel labels on objects that are more variable in shape (Perry et al., 2010).

**Language as an embodied system**

A dynamic systems theory perspective of development views the brain-behavior relationship as nested and embodied. In this view, the nervous system is nested in the body, which is then nested in the environment. Importantly, the relationships between each level are bidirectional, such that the brain can impact the body, and the body can impact the brain (Thelen, 2000). Many studies on infant development support the bidirectional relationship between the brain and behavior.

Returning to the A-not-B task, there is evidence that the decisions infants make when selecting a target location involves cognition but is also an embodied act. Several studies have demonstrated the embodiment of the A-not-B task. Smith and colleagues (1999) simply shifted the posture of the infants from sitting during the A trials to standing during the B trials. With this bodily shift, infants did not perseverate during the B trials. Similarly, another iteration of this task altered the bodily experience during the task, by having infants perform the A trials with
small weights on their arms, or with no weights, and then perform the B trials without weights or with weights if they had not worn them previously (Clearfield et al., 2006). Changing the feel of the arm between A and B trials, led to a reduction in rates of perseveration. These studies have shown a strong connection between the brain and body, suggesting the embodied nature of cognition. Notably, cognition is embodied and influenced by bodily sensations and movements.

The coupling of brain and body can also be seen in infant word learning. Smith and Gasser (2005) adapted a task from Baldwin (1993) that tested children’s ability to map novel labels onto objects when the direct object-label referent was absent. Instead, two novel objects were presented on a tabletop and then either hidden in one of two baskets in the same spatial location (Baldwin, 1993) or removed from the tabletop entirely (Smith & Gasser, 2005). Regardless of the manipulation, the unseen object was referred to by its label while directing infants’ attention to the spatial location the object was in previously. Then, the objects were brought back to their original location, and infants were asked to get one of the two objects. Infants do generally well in this task, suggesting they use spatial memory to tie an unseen object with a heard label. A dynamic field model using the same task contexts found this result highly reliable (Samuelson et al., 2017). Therefore, children can use spatial location to assist in an ambiguous labeling situation.

Additionally, to further test the embodiment of this word-learning task, Samuelson et al. (2017) switched the body posture of toddlers between seeing the objects and hearing the label. These toddlers did not successfully map the spatial location to the location, indicating that toddlers in this task use a body-centered frame of reference to figure out the ambiguous naming situation. Therefore, the embodiment during the task was crucial to learning the label-object pairing.
The previously described results of Yu and Smith (2012) also demonstrate the embodiment inherent during word learning. In this case, the body helped create moments of ideal attention toward the object. Due to the naturally shorter arms and smaller torsos infants have, when they hold an object, it physically dominates a greater proportion of their visual space than if a caregiver (with long arms and a taller torso) were to hold the same object. From the infants’ view, the toy obscures other potential referents, creating clean, visual moments for labeling. This embodied attention directly impacted the quality of learning through the self-creation of optimal moments of visual attention.

**Present Study**

The previously discussed body of literature points to a fundamental link between motor and language development and the dynamic nature of language development. Language learning appears to be incredibly dynamic such that changes in the motor system are related to increased environmental exploration, increased social bids, changing visual fields, and increased object exploration. All of these changes cascade into language development. Additionally, the processes during moments of word learning are highly dynamic, whereas moments of object manipulation during labeling relate to better word learning. However, whether object manipulation itself relates to word learning and the underlying processes of how object manipulation it might contribute to this learning are unclear. The present study aimed to further address if object manipulation impacts learning and to fill this gap by examining if the specific role of multiple, varied object views during word learning. Two experiments examined the role of multiple object views during novel word learning by focusing on whether merely experiencing multiple object views relates to similar levels of novel word learning compared to actively creating multiple object views through object manipulation. The research questions were as follows:
1. Do varied, dynamic object views contribute to in-the-moment word learning?
2. Do varied dynamic object views need to be self-generated through an infant’s object manipulation to impact learning?

Two experiments were designed to address these questions and assess whether varied object views related to novel word learning. The first experiment used novel physical objects experienced only visually each paired with a novel label. The second experiment utilized the same novel physical objects as experiment one, but the infant could manipulate them to create multiple object views while hearing a novel label.

In experiment one, two within-subject conditions differed in how the object was presented. In one condition, the object moved in a way that did not create any new or varied views for the infant. Thus, infants experienced only one, consistent view of the object. However, the object still rotated along a fixed axis, so infants still saw object rotations and movements. The second condition had the same labeling experience but differed because the object rotated, creating multiple views. Thus, the infants saw multiple, varied views of the novel object. The primary research question for this experiment was: How do varied object views that infants see impact novel word learning?

If multiple, varied views are beneficial to early word learning, then we would expect to see greater learning of label-object pairing in the condition where infants saw the objects move in a way that created varied views compared to the condition in which they saw movement but always experienced the same view of the object. However, it could be that both conditions yield low rates of learning because the task contexts their offer do not allow for manual exploration of the objects. Finally, the age range used for this experiment is an ideal time to examine early word
learning because infants are fairly adept word learners. Therefore, it could be that infants demonstrate learning in both conditions.

In the second experiment, two similar within-subject conditions were used to assess whether the types of views experienced would matter for word learning, except that in this experiment, infants manipulated objects in both conditions while hearing a novel label. In one condition, when the object was rotated along one axis, it provided a consistent view. In the other condition, when the object was rotated along another axis, it provided varying views of the object from different angles. Therefore, the experience of viewing the objects was the same as in Experiment 1, but in this study, infants were the active agents on their environment and created the views on their own. The primary research question for this second experiment was as follows: How do varied object views that infants create through their manual exploration of novel objects impact novel word learning?

If word learning benefits from object manipulations that create varied object views, we would expect infants to display higher object-label pairing learning in the condition where the child would create varied object views while interacting with the object. Alternatively, object manipulation may have an overarching benefit beyond creating varied views. In this case, there would be no difference in learning the object-label pairs in either of these conditions, with infants learning equally well from both object presentations.

Finally, we aim to compare the differences between object-word learning in both experiments as a between-subjects factor. There could be a trend that infants from both experiments learn in the varied view conditions but not in the similar view conditions. In this case, the role of multiple views would be of heightened importance over manual exploration alone. However, if only the varied view condition in Experiment 2 results in greater learning,
then the combined, synchronous experience of manipulating objects while seeing multiple views and hearing a novel label would highlight the role of dynamic, manual exploration in novel word learning.
Chapter 2: Methods—Experiment 1

Participants

Eight infants aged 20-22 months old (three males, five females) participated in this study. Participants were recruited from a database of families that had previously expressed interest in participating in research at the University of Tennessee. Three additional participants came into the lab but did not complete the task either due to becoming distracted (n = 1) or for general fussiness (n=2). The sample was homogenous and highly educated, see Table 1. Of the eight infants, seven were white, and one was of mixed ethnicity.

To determine the sample size, an effect size from a similar paradigm using same-age infants with two between subjects conditions (Booth et al., 2005) was used for the power analysis. For this study, I aimed to detect a relevant effect size with a Cohen’s $d$ of 0.8. A power analysis was performed using the two-tailed student’s t-test, with an alpha of 0.05 and a power of 0.8. From this analysis, it was found that 16 participants would be required. However, the sample size was reduced due to ongoing data collection delays due to the COVID-19 pandemic.

Materials

Infants sat in their caregiver’s lap in front of a theater. The theater was solid black and designed to obscure the view of the majority of the experimental area. The experimenter sat behind the theater, hiding from the infant’s view. Two cameras were strategically positioned around the infant to capture different angles of the study, including one camera recording the infant's behavior from the side and one recording their face and hands. The theater had an opening that was 31 cm wide and 31 cm tall. Within the theater's opening was an apparatus designed to hold the presented objects in a fixed position. The apparatus spanned the length of the opening and held the objects 20 cm high, approximately in the middle of the theater opening.
Although the theater opening was within the infants’ reaching space, they could not interact with the objects. The opening was designed to accommodate a clear plexiglass sheet, allowing the infants to see through it without being able to interact with the presented objects. Figure 1A shows the theater setup for learning trials, and figure 1B shows the setup for test trials.

The study utilized six objects, two of which were familiar: a shoe and a duck (Figure 2), while the remaining four were novel and designed using Tinkercad, an open-source 3D design platform (https://www.tinkercad.com/). Each object was specifically created to fit into both the fixed view and the varied view experimental conditions. In the fixed view condition, rotating the object along a fixed axis resulted in consistent views of the object. In contrast, the varied view condition yielded new views of the object upon rotation. Figure 3 presents each object's fixed view and varied view orientation.

All objects were 3D printed using Acrylonitrile Butadiene Styrene filament and were approximately the same size, measuring 7.62 cm in length and 5.08 cm in width at their thickest point. A hole was present in each object to mount the objects onto a rod, resulting in two sets of novel objects being printed: one set had the hole running along its horizontal axis. The other set had the hole along the vertical axis. This design allowed the objects to belong in either of the experimental conditions. The familiar objects were modified similarly so that there was a hole running through the object. However, the familiar object was always presented in its canonical orientation (e.g., the duck’s head was always oriented towards the infant and upright).

Each novel object was assigned a novel label (modi, dobu, nilla, and teema). These novel word labels had similar phonotactic probabilities and were good exemplars of potential English words ($M = 1.13$; nilla = 1.15; modi = 1.17; dobu = 1.14; teema = 1.12). The phonotactic probabilities were calculated using the Phonotactic Probability Calculator (Vitevitch & Luce,
Two types of auditory strings were used. The first type was a series of trials introducing a familiar stimulus (e.g., "There’s a duck! It’s a duck! Look at the duck! Duck!"). The second type contained the novel label and repeated the label four times within the string for the learning trials. (e.g., "Look at the Nilla! There’s a nilla! It’s a nilla! Nilla!"). Figure 4 shows the timing of the learning trials. Test trials consisted of an auditory string that prompted the infant to look toward the target stimulus and repeated the target label twice (e.g., "Where’s the nilla? Nilla! Do you see it?"). The target label always started 2000ms into the auditory string after 1500ms of silence, see figure 5.

Two questionnaires were used to collect demographic information from the caregivers and self-report measures of walking and talking onsets. In addition, the Level 2, short-form MacArthur-Bates Communicative Development Inventory (MCDI) was also used to assess the productive vocabulary size of the infants (Fenson et al., 2014). The inventory consisted of 100 words, and caregivers were instructed to mark the words their infants could independently produce. The productive vocabulary scores ranged from 4 to 51 ($M = 32.13$, $SD = 17.90$). The percentile ranks, normalized for age and gender, ranged from 5 to 60.

**Procedure**

All participants came to the lab for a one-time visit lasting approximately 30 to 45 minutes. After obtaining consent from the caregivers, infants were positioned in their caregiver’s lap in front of the theater. The caregivers were instructed to hold their infant against their chest to keep their infant steady but not restrict them tightly. Additionally, parents were instructed not to point toward the objects or say any of the object names. Finally, an experimenter behind the theater checked the positioning of the front-facing camera to ensure the infant’s face was in view. After this check, the task could begin.
This study comprised two phases within the same session with no break. First was the learning phase, where infants were introduced to four novel object-label pairs, two per condition. The learning phase consisted of two within-subject conditions, a fixed view, and a varied view condition. The order of the learning trials was pseudorandomized into four blocks, where each object-label pair was presented once. This resulted in a total of four learning trials per object-label pair. Additionally, there were six trials presenting a familiar stimulus so that the infants were cued into the word-learning task. There were two initial familiar stimulus trials to start the task and one or two between each block. In total, infants proceeded through twenty-two learning trials. An example set of the learning trial order can be seen in Table 2. Following the learning trials, infants immediately proceeded into a test phase where they were tested on their learning of the object-label pairs.

The test trials immediately followed the learning trials. Similar to the structure of the learning trials, the test trials started with two initial familiar word trials, prompting the infant to look at the target familiar object out of a pair of familiar objects. The test trials for the novel labels were pseudorandomized into four blocks, where each label was tested once within the block. Before each block, there was another familiar label test trial. An example set of the test trial order can be seen in Table 3.

**Familiar objects**

The presentation of the familiar objects was similar to the presentation of the novel objects. The trial began with the experimenter opening the curtain to reveal the object affixed to the apparatus. When the curtain was lifted, and the object was fully exposed, a secondary experimenter triggered the start of the auditory stimulus. While the auditory stimulus was playing, the experimenter moved the object, rotating it along the apparatus.
Only the experimenter’s hand was visible as the object was moved. Once the auditory stimulus ended, the experimenter lowered the curtain and removed the object to begin the next trial.

**Fixed view condition**

In the fixed view condition, infants observed the novel objects moving and rotating but without producing new or varied views of them. The trial began with the experimenter opening the curtain to reveal the novel object affixed to the apparatus. When the curtain was lifted with the object fully exposed, a secondary experimenter triggered the start of the auditory stimulus. While the auditory stimulus was playing, the experimenter rotated the object along the apparatus, with only their hand visible to the infants. After the auditory stimulus ended, the curtain was lowered, and the object was hidden from view, ending the trial.

**Varied view condition**

In the varied view condition, infants observed the novel objects moving in a way that created new or varied views of the objects. The trial began with the experimenter opening the curtain to reveal the novel object affixed to the apparatus. When the curtain was lifted with the object fully exposed, a secondary experimenter triggered the start of the auditory stimulus. During the bout of the auditory stimulus, the experimenter rotated the object along the apparatus, with only their hand visible to the infants. After the auditory stimulus ended, the curtain was lowered, and the object was hidden from view, ending the trial.

**Test Trials**

The test trials began after the learning trials without a break. The test trials were consistent regardless of the object presentation condition. Across the test phase, each object-label pair was tested four times. The target object was always tested against the other object from the same condition, and the side of the presentation of the target object was counterbalanced. For
example, label one from the varied view condition always was tested against label two from the varied view condition.

The trial began with the experimenter opening the curtain to reveal the two novel objects, which were affixed to the apparatus approximately 8cm apart. When the curtain was lifted with the object fully exposed, a secondary experimenter triggered the start of the auditory stimulus. For the test trials, the objects remained stationary. After the auditory stimulus ended, the curtain was lowered, and the objects were hidden from view, ending the trial.

**Coding**

**Learning trials**

The data from the front-facing webcam for the learning trials were hand-coded by a primary coder offline using DataVyu (Datavyu Team, Databrary Project, New York University). The learning trials were coded for whether the infant was attending to the scene or was off task. While coding, the coder could not see the object or hear when the label was provided. Therefore, the coder was blind to the condition that they were coding. Since the coder could not see the object, being off task during the learning trials was defined as the infant looking away from the center of the theater, therefore, away from the novel object presentation. While the learning trial auditory stimuli length was consistent, the timing was not always exact due to human error while facilitating physical object presentations. Therefore, a proportion of time infants spent looking at the scene was used instead of a raw measure of time.

**Test trials**

The test trials were coded using Peyecoder (Olson et al., 2020). While coding, the coder did not use sound, although the sound was available if the coder needed to check what the parent said during the task. Additionally, the coder could not see the objects that were being tested.
Therefore, the coder was unaware of the condition being tested or which side the target object was on. The videos were coded frame by frame, coding eye movement. There were four primary codes used to indicate eye movement:

1. Left, with the infant looking towards the left of the screen,
2. Right, with the infant looking towards the right of the screen,
3. Away, with the infant looking elsewhere in the room,
4. Off, when the infant began to shift their eyes from one location to another.

Importantly, “Off” was only used when the infant was transitioned from one area of interest to another (e.g., from left to right) and not for eye movements that indicated scanning of an object (e.g., up and down while staying on the left side).

**Screening of trials**

Using an exported file from Peyecoder, which indicated where the infant was looking during each frame, we could use a custom-made R script to clean the data. A trial was excluded from the analyses for the following reasons:

1. If the infant was off task for 15 or more consecutive frames.
2. If the infant did not look at the target or distractor object before the onset of the label.

During a pre-screen, trials were removed if the infant’s eyes were not visible (e.g., their arms were in front of their face). There was an additional exclusion criterion of removing the trials if the caregiver talked to the infant during label onset. However, this did not occur. Following these criteria, 28% of the trials were removed in the first window and 33% in the second window.

**Reliability**

A secondary coder coded a 20% overlap of the data. Both the learning trials and the test trials were coded for reliability. For looking during the learning trials, we reached 91.09%
reliability for the duration of looking. For looking during the test trials, we reached 88\% reliability for frame agreement and 86\% for shift agreement.

**Dependent Measure**

The mean accuracy was the primary dependent measure for this task. We used measures similar to a looking while listening task (Fernald et al., 2008), but due to using physical object presentations, which may make the task more difficult, we modified the window of analysis to include two windows after each word onset. Additionally, Fernald et al., (2008) show that infants will display only rapid looks to the target object within the first 2-3 seconds and then shift away and not return during the trial if they are very familiar with the word. We expect that there may be a delay in selecting the target object for the novel words, such that we may only see indications of learning in the second window of analysis due to the time it may take for infants to process the stimuli. To account for these potential differences in visual selection between familiar and novel words, we assessed accuracy as the proportion of time infants spent looking at the target object during the two critical windows. The first critical window was 1700ms long and started 300ms after the first label onset and ended at 2000ms following label onset. The second critical window was also 1700ms long and started 300ms after the second label onset and ended 2000ms following label onset See Figure 5 for a timeline of the trial. All codable trials were used to calculate the mean accuracy.

**Analyses**

The following factors were assessed in the present experiment: the looking at the target object during learning trials and the mean accuracy during the test trials. The dependent variable was the mean accuracy, while the condition was the independent variable. Mean accuracy was
measured as the proportion of time spent looking at the target object during each critical window. All analyses were run using SPSS.

Due to the small sample size, the data were non-normally distributed. To account for the non-normal distribution, we used non-parametric statistics for the analyses. Friedman tests, or Wilcoxon Signed Ranks tests, were used to detect differences between the within-subject factors, including a preference for objects and the proportion of time spent looking at the scene during learning trials. To assess the overall accuracy of the study, we ran a Generalized Linear Mixed Model (GLMM). This type of analysis is particularly useful when dealing with clustered and correlated data sets and non-normally distributed data. The linear model distribution was used, with the infant participant ID acting as a random factor to account for repeated measures. The mean accuracy was entered into the model as the dependent factor. The within-subject condition and the window of analysis were used as the fixed effect factors. A sequential Bonferroni was used to correct for multiple comparisons for all pairwise comparisons.

To test whether the mean accuracies were above the chance level (50%), a Wilcoxon Signed-Ranks test was used. Additionally, we ran a series of Spearman correlations to assess the impact of different participant factors on accuracy in the task, including age of onset of walking and talking, vocabulary size as assessed by the MCDI, and age. Finally, a Spearman correlation was also used to understand the relationship between the individual mean proportion of looking at the scene during the learning trials and mean accuracy.
Chapter 3: Results—Experiment 1

Looking during the learning trials

Infants spent over 86% of the learning trials on average attending to the scene (Familiar: $M = 91.29\%, SD = .07$; Fixed: $M = 86.33\%, SD = .07$; Varied: $M = 90.27\%, SD = .06$), see figure 6. In addition, a Friedman test revealed no significant differences between the familiar words and the two novel word conditions ($\chi^2(2) = -1.40, p = .161$). Therefore, infants attended to the scene during the learning trials similarly between the familiar and novel words in both conditions.

Familiar Words

A Wilcoxon Signed ranks test revealed no significant difference in mean accuracy between the two familiar objects ($Z = -8.45, p = .398$). Therefore, there was no difference between the two familiar objects, indicating no inherent preference for one object.

Mean Accuracy

The mean accuracy was measured between two windows of analysis. A Wilcoxon Signed Ranks test revealed a significant difference in mean accuracy between the two windows of analysis ($Z = -2.38, p = .017$). Mean accuracy was higher in the first window of analysis ($M = .76, SD = .09$) than in the second window ($M = .55, SD = .21$). See figure 7. Mean accuracy was significantly above the chance level in the first window of analysis ($Z = 36, SE = 7.14, p = .012$), however during the second window of analysis, mean accuracy was at chance ($Z = 23, SE = 7.14, p = .48$). Figure 8 shows this trend in the pattern of looking across the average trial. There is an initial 2000ms baseline window. During this time, infants looked to the target and distractor objects similarly. After the onset of the first label, there was a dramatic shift in looking toward the target object. This started to dip about 1600ms following label onset. During the second window of analysis, there was a slight increase in looking toward the target object, but there was
not as strong of a shift in looking at the target object. This pattern could have emerged because infants were highly reliable at selecting the target object during the first window and became bored and increasingly off-task as the trial continued.

Another measure of looking response is displayed in Figure 9, an onset contingent plot. This plot charts two courses of looking dependent on which object the infant was looking at during label onset. If they were initially looking at the distractor object during label onset, they should shift their gaze to look at the target object if they know the target label-object pair. If they were on the target object at the onset of the label, they should inhibit shifting their gaze if they know the accurate label. The onset contingent plot shows a clear indication that infants knew the familiar labels. During the first window of analysis (Figure 9A), there are proportionally greater shifts to the target object if they were on the distractor object at the onset of the label. If they were on the target object at the onset of the label, there is a lower proportion of shifting their gaze away from the target object. During the second window of analysis (Figure 9B), there was a slight increase in shifting to the target if they were on the distractor, but there was also greater shifting off the target if they were at the target during label onset. This could indicate that the infants were bored during this point of the trial since they showed a strong response at the start of the trial. These results indicate that infants could readily identify a target object among familiar words.

**Novel Words**

A series of Wilcoxon Signed Ranks tests were run to see if any difference in mean accuracy was dependent on the object, indicating object preference. Each of the four objects was compared against each other within each condition (e.g., accuracy for the modi in the fixed condition was compared against the dobu, teema, and nilla in the fixed condition). There was no
significant difference in accuracy between the objects ($ps > .180$), indicating no inherent preference or difference in learning based on the object alone.

**Mean Accuracy**

A GLMM revealed no significant effect of the window of analysis ($F (1, 28) = .032, p = .858$). Therefore, for novel words, the mean accuracy was similar in each window of analysis (First: $M = .54, SE = .044$; Second: $M = .55, SE = .044$), see figure 10. There was also no main effect of condition ($F (1, 28) = .006, p = .938$). Mean accuracy was not significantly different between the fixed condition ($M = .55, SE = .044$) and the varied condition ($M = .55, SE = .044$), see figure 11.

There was not a significant interaction between the window of analysis and the condition ($F (1, 28) = .468, p = .50$). However, the mean accuracy during both the fixed and varied conditions in Figure 12 showed a distinct pattern emerging with how much infants looked to the target object in the first and second windows of analysis. An exploratory analysis revealed that for the fixed condition for the fixed condition, infants looked at the target object slightly more in the first window compared to the second window (First: $M = .56, SE = .061$; Second: $M = .53, SE = .061$). The mean accuracies for the fixed condition in both the first and second window of analysis were at chance level (First: $Z = 26.00, SE = 7.14, p = .263$; Second: $Z = 19.00, SE = 7.14, p = .889$). For the varied condition, infants had higher mean accuracies in the second window ($M = .58, SE = .061$) than in the first window of analysis ($M = .52, SE = .061$). While not significant, there is a trending effect that infants’ mean accuracy was above the chance level in the second window of analysis in the varied condition ($Z = 30, SE = 7.14, p = .09$). Mean accuracy was at the chance level during the first window of analysis ($Z = 27, SE = 7.14, p = .40$).
Figure 13A shows the time course of looking for the fixed condition. During the 2000ms baseline window, infants looked to both the target and distractor object. After the onset of the first label, infants looked at the target and distractor object equally during the first ~1200 ms of the window of analysis. Then they shifted looking towards the target object until the end of the analysis window. During the second window of analysis, there was a slight shift in looking at the target object, but infants mostly looked between the distractor and target at similar rates. Figure 13B shows the time course of looking across the trial for the varied view condition. During the baseline window, infants initially looked at both objects and then showed a slight preference for the target object. In the first window of analysis, looking at the target decreased from the baseline but had a delayed rise outside the window of analysis. In the second window of analysis, there seemed to be a learning response, with infants demonstrating an increase in the proportion of looking at the target object.

Figure 14 A and B shows the onset contingent plot for the fixed condition. There is a similar delay in a learning response seen here. In the later part of the first window of analysis (Figure 14A), infants shift to the target object proportionally more if they were initially on the distractor object. Similarly, if they were on the target object, they would have proportionally fewer shifts off the target object. Half of this shift occurs outside the window of analysis, so it is possible that infants had a delayed response to the first label. However, in the second window of analysis (Figure 14B), there is no indication of learning. Infants tended to only shift from the distractor to the target 50% of the time. Similarly, if they were on the target object at label onset, they shifted off the target about 50% of the time. Together, these findings demonstrate a delayed response in selecting the target object. There is a distinct pattern to their responses that show some evidence of learning in the varied condition especially.
Variability

Infants displayed considerable variability in their learning (see Figure 15). Five infants had higher mean accuracies in the varied than the fixed condition. Three infants had higher mean accuracies in the fixed condition than the varied condition. Overall, there is a trend that infants are doing better in the varied condition than the fixed. However, for three of the infants that did better in varied than fixed, they had below-average accuracy in the fixed view condition. Therefore, this trend could come down to regression to the mean because these three infants only hit around 50% in the varied condition, which would be at chance level.

Correlations between accuracy and participant factors

Familiar words

A series of Spearman correlations were run on the average accuracy between the two windows of analysis and MCDI, age, the onset of walking, and the onset of talking. Two participants were excluded from correlations involving talking onset because caregivers indicated a severely low age for the onset of talking (4 months old). There was not a significant correlation for the age-normed MCDI percentile rank ($r(8) = .07, p = .866$), age ($r(8) = .49, p = .22$), walking onset ($r(8) = 0, p = 1$), or talking onset ($r(6) = -.53, p = .28$). See figure 16 for plots. Additionally, a Spearman correlation was used to see if there was a relationship between the proportion of time infants spent on task during the learning trials and mean accuracy for the familiar words. There was not a significant trend between these two factors ($r(8) = -.24, p = .568$)

Novel words

To evaluate the correlation between participant factors, including MCDI, age, age of onset of walking, age of onset of talking, and mean accuracy in both novel word conditions, a
A series of Spearman correlations were performed. In the fixed condition, there was no significant relationship between mean accuracy and infants’ MCDI percentile rank \((r(8) = .20, p = .629)\). There was no trend for age \((r(8) = .19, p = .65)\), walking onset \((r(8) = .96, p = .82)\), or talking onset \((r(6) = .71, p = .117)\). Similarly, for the varied condition, there was no significant relationship between mean accuracy and infants’ MCDI percentile rank \((r(8) = -.23, p = .312)\). There was no trend in walking onset \((r(8) = .036, p = .932)\) or talking onset \((r(6) = .706, p = .117)\). There was, however, a trending relationship between age and accuracy, where older infants tended to have lower mean accuracies \((r(8) = -.702, p = .052)\). See figure 17 for plots for the fixed condition and figure 18 for the varied condition. These results show that there is no evidence that the recorded individual differences influence variability in mean accuracies.

Finally, a Spearman correlation was used to see if there was a relationship between the proportion of time infants spent on task during the learning trials and mean accuracy in the fixed and varied conditions. There was not a significant trend between looking during the learning trials and mean accuracy in either the fixed \((r(8) = .14, p = .736)\) or varied conditions \((r(8) = -.38, p = .352)\); see figure 19.
Chapter 4: Summary—Experiment 1

The aim of this experiment was to understand the difference in word learning between fixed and varied views when infants passively observed those during a novel-object labeling. Although there was no overall difference in mean accuracy between the fixed and varied conditions, several distinct patterns emerged. First, there was a trending increase in looking at the target object in the fixed view conditions. However, this occurred late in the first window of analysis and did not peak until the window was over (2000ms following label onset). A similar delay was seen with the onset contingent plot, showing that infants shifted off the distractor object to the target object in the later half of the first window of analysis, with a peak in shifting just outside the window.

Similarly, in the varied condition, there was a slight peak in looking at the target object outside the first window of analysis. During the second window of analysis, there was a pattern of looking toward the target, which resulted in a trend in the mean accuracies for the second window to be above the chance level. The onset contingent plot showed a similar delay in the varied condition compared to the fixed condition. There is a peak in the proportion of shifts to the target object from the distractor object. However, this occurs just outside the first window of analysis. There is a stronger learning response in the second window of analysis, with higher shifts from the distractor to the target and lower shifts away from the target. These trends suggest that infants learned the words in the varied condition.

There was also no relationship between vocabulary size and mean accuracy in either condition. This was surprising because it could be that infants with larger vocabulary size would do better in a hypothesized challenging condition, such as the fixed view condition. Similarly, there was no relationship between walking onset and accuracy. Previous work has shown that
earlier onsets of walking are related to increased vocabulary sizes (Walle & Campos, 2014), but there does not seem to be an impact on in-the-moment word learning. There was a trending effect of a negative correlation between age and accuracy in the varied condition. Infants that were younger tended to have higher mean accuracies than older infants. This was an unexpected finding because, overall, there might be the expectation that older infants are more accurate. Even so, this would probably be related to vocabulary size, which did not impact accuracy. Three infants were 88 weeks old that ranged from 50-70% mean accuracy, and then the older infants ranged from 90 to 94 weeks old with slightly lower mean accuracies. It could be due to the sample size, with the sample skewing slightly younger within the 20-to-22-month-old range, resulting in this trending correlation.

This experiment had no object manipulation and, because of this, was hypothesized to have lower rates of learning. Importantly, neither window of analysis resulted in infants being above chance level in the fixed condition. Although there are some indications of learning, it is not strong. Interestingly, there are similar delays in selecting the correct object in both conditions. This could indicate that the task was difficult for the participants, and it took longer for them to fixate on the target object. Infants mean accuracies were high for the familiar words, which suggests they were capable of selecting target objects. However, the novel words may have been a more difficult learning environment for the infants. Finally, while not significant, there was a trending effect that the second window of analysis for the varied condition was above the chance level. This condition was hypothesized to have better learning rates due to infants experiencing multiple, varied views of the novel object. It could be in a challenging learning environment that only allowed for visual exploration that the varied views boosted learning rates.
Overall, it is unclear whether varied views impacted learning in this experiment. However, there seems to be a slight trend that the varied views improved learning. The second experiment allowed for a better understanding of the role of multiple, varied object views by allowing infants to explore the objects manually. Thus, infants do not passively observe fixed or varied views in the second experiment. Instead, they create these differing views through their own object manipulation. Adding in the factor of object manipulation will allow for a better understanding of how varied views impact novel word learning.
Participants

This study involved eight infants between 20 and 22 months old (M = 20.25, 4 of whom were female). Five additional participants were brought into the lab. One participant finished the task, but the test trials were not recorded due to a technical error. In addition, one participant was excluded for being off task during the test trials. Finally, three participants did not finish the task due to general fussiness. The participants were recruited from a database of families who had previously shown interest in research. The sample was relatively homogenous and well-educated, as shown in Table 1.

A power analysis was conducted using an effect size from a similar study by Booth et al. (2005), which used the same age group of infants and two between-subjects conditions to determine the appropriate sample size. The goal was to detect a relevant effect size with a Cohen's d of 0.8. The power analysis used a two-tailed Student's t-test with an alpha of 0.05 and a power of 0.8, indicating that 16 participants would be necessary. However, due to ongoing data collection delays caused by the COVID-19 pandemic, the sample size had to be reduced.

Materials

The theater was intentionally designed to be solid black to obscure the view of the rest of the experimental area and the experimenter from the infant. In addition, two cameras were strategically positioned around the infant to capture different angles of the study. One camera recorded the infant's manual exploration behavior from the side, while the other recorded their face and hands. Figure 1 shows the experimental set-up.

The study used the same 3D printed objects as the first experiment and used the same four novel labels (modi, dobu, nibla, and teema). Two types of auditory strings were used: the
first introduced a familiar stimulus, while the second contained the novel label and repeated it four times within the string. Figure 4 illustrates the timeline for the learning trial auditory stimuli. Test trials consisted of an auditory string that prompted the infant to look toward the target stimulus and repeated the target label twice. The target label always began 2000ms into the auditory string, following 1500ms of silence (see Figure 5).

Two questionnaires were administered to collect demographic information from the caregivers and self-report measures of walking and talking onsets. Additionally, the Level 2, short-form MacArthur-Bates Communicative Development Inventory (MCDI) was used to assess the productive vocabulary size of the infants (Fenson et al., 2014). The inventory consisted of 100 words, and caregivers were instructed to mark the words their infants could independently produce. The productive vocabulary scores ranged from 2 to 65 ($M = 30.00, SD = 26.15$). The percentile ranks, normalized for age and gender, ranged from 5 to 90.

**Procedure**

The study consisted of two phases conducted in a single session without interruption. The first phase was the learning phase, in which the infants were introduced to four novel object-label pairs. This was immediately followed by the test phase, where the infants were tested on their learning of the object-label pairs. The learning phase included two within-subject conditions, the fixed view and varied view conditions. The order of the learning trials was pseudorandomized into four blocks. Each object-label pair was presented once within the block, resulting in four learning trials per object-label pair. Additionally, there were six trials presenting a familiar stimulus to cue the infants into the word-learning task. These trials were distributed throughout the learning phase, with two initial familiar stimulus trials and one or two between each block, resulting in twenty-two learning trials. Table 2 shows an example set of learning trial orders.
The test trials immediately followed the learning trials. They began with two initial familiar word trials, prompting the infant to look at the familiar object out of a pair of familiar objects. The test trials for the novel labels were pseudorandomized into three blocks, with each label being tested once within the block. Before the second and third blocks, there was another familiar label test trial. Table 3 shows an example set of test trial orders.

**Familiar objects**

The procedure for presenting familiar objects was similar to that of novel objects. Each trial began with the experimenter pulling back the curtain to reveal the object fixed on the apparatus within the infant's reach. A secondary experimenter initiated the auditory stimulus when the object was fully exposed. During the stimulus presentation, infants could manipulate the object on the apparatus. After the auditory stimulus ended, the experimenter lowered the curtain and removed the object to prepare for the next trial.

**Fixed view condition**

In this condition, infants could manipulate the novel object and create object motion. However, the view that infants created did not change. Each trial began with the experimenter pulling back the curtain to reveal the object fixed on the apparatus within the infant's reach. A secondary experimenter initiated the auditory stimulus when the object was fully exposed. During the stimulus presentation, infants could manipulate the object on the apparatus. After the auditory stimulus ended, the experimenter lowered the curtain and removed the object to prepare for the next trial.

**Varied view condition**

In this condition, infants could manipulate objects and create object motion that created multiple views of the object. Each trial began with the experimenter pulling back the curtain to
reveal the object fixed on the apparatus within the infant's reach. A secondary experimenter initiated the auditory stimulus when the object was fully exposed. During the stimulus presentation, infants could manipulate the object on the apparatus. After the auditory stimulus ended, the experimenter lowered the curtain and removed the object to prepare for the next trial.

**Test Trials**

The test trials began after the learning trials. Before the test trials began, a sheet of plexiglass was slid into the front of the theater, so infants could not interact with the objects during the test trials. The test trials were uniform across all object presentation conditions, with each object-label pair being tested four times. The target object was tested against the other object from the same condition, with the side of the target object presentation being counterbalanced. For instance, label one from the consistent view condition was always tested against label two from the consistent view condition. The trial commenced with both objects displayed on the screen, out of the infant's reach, in a single view. The auditory stimulus coincided with the start of the trial, prompting the infant to look towards the target object, and concluded after the auditory stimulus was played. Each test trial lasted about 7,000ms.

**Coding**

**Learning trials**

During the learning trials, the data captured by the front-facing webcam was manually coded offline by a primary coder using DataVyu (Datavyu Team, Databrary Project, New York University). The coder evaluated whether the infant was oriented towards the scene or off-task, i.e., oriented elsewhere. It is important to note that the coder was unaware of the specific condition or object being tested as they did not have access to the object or accompanying label.
Off-task behavior was defined as the infant looking away from the center of the theater, away from the presentation of the novel object.

Manual exploration of the object was also coded by observing whether the infant's hands were on the object. This was performed using the side camera that captured both the object and the infant's hands. One participant was excluded from object manipulation analyses due to a recording issue resulting in no side camera data.

**Test trials**

Eye movements were coded frame by frame for the test trials using Peyecoder (Olson et al., 2020). The primary coder had access to sound but only used it, if necessary, to check what the parent was saying during the task. In addition, the coder was blinded to the condition being tested and which side the target object was on as they could not see the objects. Four primary codes were used to indicate eye movements:

1. Left, with the infant looking towards the left of the screen.
2. Right, with the infant looking towards the right of the screen.
3. Away, with the infant looking elsewhere in the room.
4. Off, used when the infant transitioned from one area of interest to another (e.g., from left to right) and not for scanning eye movements (e.g., up and down while staying on the left side).

**Screening of trials**

Using a customized R script, we processed the exported file from Peyecoder to obtain clean data on the infant's gaze direction in each frame. We excluded trials from analysis if the infant was off task for 15 or more consecutive frames or did not look at the target or distractor object before the label onset. During a pre-screen, we also removed trials where the infant's eyes
were not visible or obstructed. There was also a criterion to remove trials where the caregiver talked to the infant during label onset, although this was not necessary. For the first analysis window, 29.27% of trials were removed, and for the second window of analysis, 30.49% of trials were removed based on these criteria.

Reliability

A secondary coder coded a 20% overlap of the data. Both the learning trials and the test trials were coded for reliability. For looking during the learning trials, we reached 94.44% reliability for the duration of looking. For manipulation, during the learning trials, we reached 96.00% reliability for the duration of the manipulation bout. For looking during the test trials, we reached 89.28% reliability for frame agreement and 91.36% for shift agreement.

Dependent Measure

The primary dependent measure for this task was the mean accuracy, assessed as the proportion of time infants spent looking at the target object during a critical window. We used measures similar to a looking while listening task (Fernald et al., 2008), but due to using physical object presentations, we modified the window of analysis to include two windows after each word onset. We assessed accuracy as the proportion of time infants spent looking at the target object during the two critical windows. The first critical window was 1700ms and started 300ms after the first label onset and ended 2000ms following label onset. The second critical window was 1700ms and started 300ms after the second label onset and ended 2000ms following label onset. See Figure 5 for a timeline of the trial. All codable trials were used to calculate the mean accuracy.
Analyses

In this experiment, we examined the effects of three factors on mean accuracy during test trials: looking at the target object during learning trials, manipulation during learning trials, and condition. Due to delays caused by the COVID-19 pandemic, the sample size for experiment two was only half of what was needed, according to the power analysis. As a result, the data did not follow a normal distribution, and non-parametric statistics were used for the analyses. To assess the overall accuracy of the study, we conducted a Generalized Linear Mixed Model (GLMM), which is suitable for clustered and correlated data sets that are non-normally distributed. The GLMM included the infant participant ID as a random factor to account for repeated measures, with the mean accuracy as the dependent factor and the within-subject condition and window of analysis treated as the fixed effects factors. We used a sequential Bonferroni correction to address multiple comparisons in all post-hoc, pairwise comparisons.

We used a Wilcoxon Signed-Ranks test to determine if mean accuracies were above the chance level (50%). We also conducted a series of Spearman correlations to examine the influence of participant factors, including the age of walking and talking onset, vocabulary size as assessed by the MCDI percentile rank, and age, on mean accuracy. Additionally, we used a Spearman correlation to investigate the relationship between task dynamics, such as the amount of manipulation during the learning trials, the proportion of time spent looking at the scene during the learning trials, and mean accuracy.
Chapter 6: Results—Study 2

Looking during the learning trials

Infants spent over 88% of the time on average during the trials attending to the scene (Familiar: $M = 92.92\%, SD = .07$; Fixed: $M = 94.5\%, SD = .04$; Varied: $M = 88.81, SD = .09$), see figure 20. In addition, a Friedman test revealed no significant differences between the familiar words and the two conditions ($\chi^2(2) =1.75, p = .415$). Therefore, infants looked at the scene similarly, regardless of whether it was a familiar or novel word and regardless of the type of object movement.

Manipulation during the learning trials

These analyses assess infants' mean proportion of time manipulating the object during the initial learning trials. There was not a significant difference in manipulation between the familiar words and the fixed and varied conditions, as shown by a Friedman test ($\chi^2(2) =2.00, p = .368$; Familiar: $M = .39, SD = .11$; Fixed: $M = .39, SD = .17$; Varied: $M = .32, SD = .23$). Thus, infants spent a similar amount of time manipulating the objects in each condition, see Figure 21.

Familiar Words

A Wilcoxon Signed ranks test revealed no significant difference in mean accuracy between the two familiar objects ($Z =-1.4, p = .161$). Therefore, there was no preference between the objects.

Mean Accuracy

The mean accuracy during the first window of analysis (300ms – 2,000ms) was significantly above the chance level, which was 50% ($Z = 32.00, SE = 7.141, p = .05; M = .66, SD = .18$). For the second window of analysis (2,800ms – 4,500ms) the mean accuracy was not above chance level ($Z = 28.00, SE = 7.141, p = .16; M = .60, SD = .24$), see Figure 22. Figure 23
shows the time course of the proportion of looking at the target object from the initial 2000ms baseline window to the two windows of analysis. This figure shows that infants looked at the distractor and target object during the baseline window. After the onset of the first label, during the first window of analysis, infants rapidly shifted to look at the target object. There is a slight decrease in looking to the target, with looking at both objects equally before looking shifted back to the target object during the second window of analysis.

Additionally, figure 24 shows the onset contingent plot. This plot shows how infants’ gaze shifted depending on where they looked at label onset. If they were initially looking at the distractor object during label onset, they should shift their gaze to look at the target object if they know the labeled object. If they were on the target object at the onset of the label, they should not shift their gaze if they know the accurate label. This figure demonstrates that during the first window of analysis (Figure 24A) if infants were on the distractor object, they show proportionally greater shifts to the target object. If infants were on the target object, they sometimes shifted away but mostly stayed on the target object during the first window of analysis. The second window shows a similar trend but a weaker trend (Figure 24B). There was a slight increase in shifting to the target object if they were on the distractor at label onset. If they were on the target object, they generally did not shift their gaze, but as the second window of analysis continued, infants shifted off the target object more.

**Novel Words**

A series of Wilcoxon Signed Ranks tests were run to see if any difference in mean accuracy was dependent on the object, indicating object preference. Each of the four objects was compared against each other within each condition (e.g., accuracy for the modi in the fixed condition was compared against the dobu, teema, and nilla in the fixed condition). There was no
significant difference in accuracy between the objects ($ps > .18$), indicating no inherent preference or difference in learning based on the object alone.

**Mean Accuracy**

A GLMM revealed no effect of window of analysis ($F = .34 \ (1, \ 28), \ p = .856$), with both the first ($M = .50, \ SE = .04$) and second windows ($M = .50, \ SE = .04$) showing similar overall accuracy, see figure 25. There was a main effect of condition ($F = 7.836 \ (1, \ 28), \ p = .009$). The mean accuracy was higher in the varied condition ($M = .56, \ SE = .04$) than in the fixed view condition ($M = .44, \ SE = .04$); see figure 26.

There was no window-by-condition interaction ($F = 1.05 \ (1, \ 28), \ p = .32$). However, there are trending differences between the conditions and windows; see figure 27. In general, the trend showed that for the varied condition, mean accuracy was higher in the second window ($M = .58, \ SE = .05$) than in the first window ($M = .54, \ SE = .05$). In fact, mean accuracy was not above chance level for the first window ($Z = 27.00, \ SE = 7.14, \ p = .575$) but for the second window, the mean accuracy was significantly above chance level ($Z = 33.00, \ SE = 7.141, \ p = .036$). For the fixed condition, there was a trend that mean accuracy was higher in the first window ($M = .46, \ SE = .05$) than in the second window ($M = .41, \ SE = .05$). However, neither of these windows had a mean accuracy above chance level (First: $Z = 14.00, \ SE = 7.14, \ p = .575$; Second: $Z = 10.00, \ SE = 7.14, \ p = .263$).

Figure 28A shows the time course of the proportion of looking at the target object from the initial 2000ms baseline window to the two windows of analysis for the fixed condition. In this figure, infants looked toward the target object slightly more than the distractor during the baseline window. After the onset of the first label, during the first window of analysis, infants rapidly shift to looking at the distractor object and demonstrate this perseveration throughout the
trial. Between the two windows, infants tended to look towards the distractor for a greater proportion of frames. During the second window of analysis, infants looked to the target object proportionally less. Figure 28B shows the time course of the proportion of looking at the target object during the varied condition. During the 2000ms baseline window, infants looked to both the target and distractor objects. After the first label onset, infants shifted their looking toward the target. During the second window of analysis, infants also demonstrated increased looking toward the target object.

Figure 29 A and B shows the onset contingent plot for the fixed condition. Figure 29A shows that during the first window of analysis if infants were on the distractor object at label onset, they did not shift to the target object. If infants were on the target object, they tended to stay on it following the first window of analysis. The second window shows a similar trend (Figure 29B), with infants infrequently shifting to the target object if they were originally fixated on the distractor object. Figure 29 C and D shows the onset contingent plot for the varied condition. In this case, the pattern of shifting looks toward the target object was similar to what was seen in the familiar condition (Figure 24). During the first window of analysis (Figure 29C), if infants were initially on the distractor object, they showed greater shifts toward the target object. This same pattern emerged in the second window of analysis (Figure 29D).

**Variability**

Infants showed considerable variability in their task performance. Figure 30 shows the mean accuracy in the fixed and varied conditions with the individual performance overlaid on the plot. Six of the eight infants had higher mean accuracies in the varied condition than the fixed. Only one infant had a lower mean accuracy in the varied condition than the fixed condition. One
infant had similar mean accuracies across both conditions. Overall, the trend was that infants had higher rates of looking at the target object in the varied condition compared to the fixed.

**Correlations between word learning and participant factors**

*Familiar words*

The mean accuracy across the two windows was averaged to compare the relationship between the accuracy and participant factors, including MCDI, age, age of onset of walking, and age of onset of talking. A Spearman correlation revealed no significant relationship between mean accuracy of the familiar words and infants’ MCDI percentile rank ($r(8) = -.13, p = .754$). There was no trend for the participant’s age ($r(8) = -.50, p = .20$), walking onset ($r(8) = -.56, p = .14$), or talking onset ($r(8) = .024, p = .955$). See figure 31 for the plots.

A Spearman correlation was used to see if the proportion of time infants spent looking at the scene during the learning trials was related to their mean accuracy during test trials. There was no significant trend between these two factors ($r(8) = -.342, p = .408$). Additionally, a Spearman correlation was used to see if there was a relationship between the proportion of time infants spent manipulating the object in the learning trials and mean accuracy for familiar words. Again, there was not a significant trend between these two factors ($r(7) = .43, p = .337$).

*Novel words*

A series of Spearman correlations were used to assess the relationship between mean accuracy in both novel word conditions and participant factors, including MCDI, age, age of onset of walking, and age of onset of talking. In the fixed condition, there was no significant relationship between mean accuracy of the familiar words and infants’ MCDI percentile rank ($r(8) = .22, p = .606$). There was no trend walking onset ($r(8) = -.49, p = .31$), or talking onset ($r(8) = -.47, p = .233$). However, for the participant’s age, there was a strong, negative
correlation with mean accuracy ($r(8) = -0.91, p = .002$); see figure 32. Similarly, for the varied condition, there was no significant relationship between mean accuracy of the familiar words and infants’ MCDI percentile rank ($r(8) = 0.21, p = .627$). There was no trend for age ($r(8) = -0.21, p = .62$), walking onset ($r(8) = -0.09, p = .83$), or talking onset ($r(8) = -0.31, p = .45$), see figure 33.

Additionally, a Spearman correlation was used to see if the proportion of time infants spent looking at the scene during the learning trials was related to their mean accuracy during test trials. There was no trend for the fixed condition ($r(8) = 0.38, p = .352$). However, a strong, positive correlation emerged for the varied condition ($r(8) = 0.71, p = .047$). Therefore, infants who spent more time looking at the scene for the varied view objects during learning trials had higher mean accuracies during the test trials, see figure 34.

Finally, a Spearman correlation was used to see if there was a relationship between the proportion of time infants spent manipulating the object in the learning trials and mean accuracy in the fixed and varied conditions. There was not a significant trend between these two factors in the fixed condition ($r(7) = 0.29, p = .53$) or the varied condition ($r(7) = 0.43, p = .34$, see figure 35).
Chapter 7: Summary—Experiment 2

The aim of this experiment was to understand the difference in word learning between fixed and varied views when infants created those views through their own object exploration during a novel-object labeling. Overall, rates of object manipulation were low, with infants’ hands on the object for only about 30% of the trial. This could be due to the artificial way the object was affixed to the apparatus. Infants also manipulated the objects for the same time between the two conditions. Despite the similarity in the proportion of time manipulating the object, learning was significantly different in the two conditions. Overall, infants had a higher mean accuracy in the varied condition compared to the fixed condition. This indicates that multiple, varied views infants create through their manual exploration benefit learning.

There was also evidence of learning based on infants’ looking patterns in the varied condition. In the first and second windows of analysis, there is a pattern of infants looking at the target object more than the distractor object. The onset contingent plots show similar patterns of evidence of learning in the varied condition and not in the fixed condition. In the first window of analysis in the varied condition, infants shift off the distractor and go to the target proportionally more. They also stay on the target if they were already fixated there during label onset. The second window of analysis shows the same pattern. If infants were on the distractor during the label onset, they would shift to the target. There is also very little shifting away from the target if they were already fixated on the target during the label onset. Taken together, the looking pattern during the varied condition indicates that infants learned the novel labels.

There is no similar pattern showing evidence of learning in the fixed condition. During both windows of analysis for the fixed condition, infants look between the target and distractor objects similarly, with some perseveration toward the distractor object. The onset contingent plot
shows that during the first window of analysis, infants shift off the target object proportionally more than shifting from the distractor to the target object. This trend continues in the second window of analysis. These trends indicated that infants did not learn the novel labels in the fixed condition.

There was no relationship between the proportion of time spent manipulating the objects during the learning trials and mean accuracy. This was surprising because more manipulation of objects would mean infants experienced more views of the object. However, infants were largely homogenous in the time spent manipulating the objects. Interestingly though, there was a significant, positive correlation between the proportion of time spent attending to the scene in the learning trials and the mean accuracy of words in the varied condition. This implies that when infants observed the varied views they created more, they had higher accuracy. There was also no relationship between mean accuracy and the different participant factors, including MCDI scores, the onset of walking, and the onset of talking. Although we see variability in how infants do in the task, it does not seem to be linked to these individual differences. There was, however, a significant, negative correlation between age and mean accuracy in the fixed condition. In this case, the older infants did worse in this condition. This is a surprising finding since there are no predictions that age would impact learning outside of other factors, such as vocabulary size.

In this experiment, it was hypothesized that the varied condition would have higher rates of learning due to infants creating multiple, varied object views through their own manual exploration of novel objects. Although there was variability in the sample, most infants showed higher mean accuracies in the varied condition compared to the fixed condition. Additionally, learning was above chance level in the second window of analysis for the varied condition but not for either window in the fixed condition. An alternative hypothesis was that infants would
learn similarly in both conditions, supporting an overall benefit of manual exploration during labeling. These results demonstrate an overall benefit to creating multiple, varied views. This suggests that the link between object manipulation and word learning comes down to what object manipulation allows infants to do: create new and varied object views. Between the two experiments, there is a trend that the varied condition resulted in higher mean accuracies. Directly comparing the results of the two experiments allowed for a better understanding of the relative contribution of the motor system to novel word learning.
Chapter 8: Results—Experiments 1 and 2

Experiments 1 and 2 utilized identical object presentations and auditory stimuli. The methodological difference for these experiments was in how the objects were presented; in Experiment 1, the experimenter moved the object, and in Experiment 2, the infant could manually explore the object. The similar methodology allowed for the experiments to be compared statistically. First, a Mann-Whitney test was used to compare the proportion of looking at the scene between experiments. Then, a GLMM was used with the experiment as a between-subject factor and the condition and window of analysis as fixed factors. The mean accuracy was the dependent factor. Finally, the infant participant ID was included as a random factor to account for repeated measures.

Looking during the learning trials

A Mann-Whitney test was used to understand if the proportion of looking at the scene during learning trials differed between the two experiments. Between experiments, infants looked at the scene during learning trials at similar proportions ($Z = -1.69, p = .090$, see figure 36). Therefore, between the two experiments, infants similarly looked at the object movement, either self-created or passively observed.

Mean accuracy

The GLMM did not show a difference in mean accuracy between the two experiments ($F(1, 56) = 1.18, p = .281$). Overall, the performance was similar in Experiment 1 ($M = .55, SE = .032$) and Experiment 2 ($M = .50, SE = .032$), see figure 37. There was no difference between overall mean accuracy in the window of analysis ($F(1, 56) = .001, p = .971$), with the first window ($M = .52, SE = .029$) showing similar levels of accuracy as the second window ($M = .52, SE = .029$), see figure 38. However, there was a marginally significant, trending effect of
condition \( (F(1, 56) = 2.98, p = .09) \). In the varied condition, mean accuracy tended to be higher than the fixed condition across experiments \((\text{Varied: } M = .56, SE = .029; \text{Fixed: } M = .49, SE = .029)\), see figure 39.

There was no experiment by condition interaction \( (F(1, 56) = 2.559, p = .115) \); see figure 40. However, there was a trending effect within the interaction that the fixed condition in experiment one \((M = .55, SE = .041)\), with no object manipulation, had a higher mean accuracy than in experiment two \((M = .44, SE = .041)\). Finally, there was no experiment by window interaction \( (F(1, 56) = .065, p = .80, \text{fig 41}) \) or a three-way interaction between experiment, condition, and window \( (F(1, 56) = .003, p = .955) \). These results indicate that infants demonstrated similar rates of learning in both experiments. However, there was a trend that infants overall had higher mean accuracies when objects had multiple varied views compared to fixed, consistent views.
Chapter 9: General Discussion

A substantial body of evidence demonstrates the link between motor and language development (He et al., 2015; Libertus & Violi, 2016; Oudgenoeg-Paz et al., 2012; Walle & Campos, 2014). Importantly, the relationship is mediated by the exploration in which infants engage when they gain access to more of their environment (Oudgenoeg-Paz et al., 2015). Previous research has suggested that manual exploration of objects during labeling moments can benefit word learning, but the mechanism underlying this effect was unclear. According to Yu and Smith (2012), manual exploration leads to optimal moments of visual attention. Holding an object allows it to dominate more of the visual field, obscuring other potential referents. DiMercurio (2019) found that infants who did not learn in a visual-only condition could learn the label when they held the object, suggesting that manual exploration had some benefit beyond creating optimal moments of visual attention. However, it was unclear whether the benefit observed was due to the haptic experience of holding the object or the fact that holding an object allows infants to generate varied views of the object.

Experiment 1 had two within-subject conditions that differed how the novel object was presented. The first condition had no new or varied views of the object, while the second condition had multiple views due to object rotation. The research question was whether passively observed varied object views impact word learning. If so, we would expect higher rates of learning in the condition with varied views. However, it was possible that there may have been lower rates of learning in both conditions due to the absence of object manipulation. Alternatively, there could have been evidence of learning in both conditions because infants in this age range are fairly adept word learners. In this experiment, there was overall mixed evidence for learning. Infants were clearly able to identify familiar targets but were not as clear
in how they identified novel targets. Experiment 2 presented physical objects to infants and had two within-subject conditions that differed in the types of views experienced. One condition had consistent views, while the other had multiple views due to object rotation. The research question was whether varied object views from manual exploration impacted word learning. If so, we would have expected higher learning rates in the condition with varied views. Alternatively, there may have been a general benefit from the haptic exploration of objects, resulting in no difference in learning between the conditions.

**Fixed object views**

Experiment 1 showed some indications of learning for the fixed condition, but it was not very strong. The results showed a clear increase in looking at the target object, which occurred late in the first window of analysis and did not peak until after the window was over, which was 2000ms following label onset. The onset contingent plot also showed that infants shifted their gaze from the distractor object to the target object in the later half of the first analysis window, with a peak in shifting just outside the window. When infants were prompted again to look at the target object before the second window of analysis, they only showed a slight indication of visually selecting the target object. Still, this response was weak and barely above equal looking between the target and the distractor.

There was even less evidence for learning in the fixed condition in the second experiment. During both windows of analysis, infants looked at the target and distractor objects similarly. The onset contingent plot also showed that infants shifted away from the target object more often than from the distractor to the target object during the first window of analysis. This trend continued in the second window of analysis, indicating that infants did not learn the novel labels in the fixed condition when they manipulated the object.
The findings suggest that the fixed view condition is not the ideal presentation style for novel word learning. When infants passively experienced the fixed views, there was an overall delay in looking toward the target object following label onset. This suggests that the target object may have been difficult to identify. When infants created fixed views independently, there was no indication of learning. Surprisingly, the evidence for learning looked slightly better, though still weak, in the first experiment compared to the second. One of the original predictions was that haptic experience might have an overall benefit during labeling moments. However, there was no clear benefit to haptic experience and object manipulation overall. Instead, there seemed to be some benefit regarding varied views.

**Varied object views**

The evidence for learning in the varied condition in the first experiment is mixed. There was a delayed peak in looking at the target object outside the first window of analysis. However, during the second window of analysis, there was a trend of looking toward the target, resulting in a trend of mean accuracies for the second window to be above the chance level. The onset contingent plot in the varied condition showed a similar delay, with a clear peak in the proportion of shifts to the target object from the distractor object just outside the first window of analysis. However, in the second window of analysis, there was a clear pattern of large proportions of shifting from the distractor object to the target object following label onset. Similarly, there were few shifts from the target object to the distractor following label onset if they were already on the target, indicating that infants knew the label.

When infants passively experienced multiple, varied views, there was a stronger learning response in the second window of analysis, with higher shifts from the distractor to the target and lower shifts away from the target. These findings suggest that infants may have had a more
difficult time mapping the label to the object in the varied condition. Still, the varied views of the novel object may have boosted their rates of learning in a challenging learning environment that allowed only for visual exploration. Furthermore, while not statistically significant, there was a trending effect in the second window of analysis for the varied condition, where the mean accuracy was above the chance level. This is interesting because the varied condition was hypothesized to result in better rates of learning, given the multiple, varied views of the novel object presented to the infants. It is possible that in a challenging learning environment, where infants were limited to visual exploration, varied views of the novel object improved their learning rates.

In the second experiment with object manipulation, however, the evidence for learning in the varied condition is strong. First, this experiment had an overall main effect on the condition, with the varied condition having higher mean accuracies than the fixed condition. The varied condition in this experiment also resulted in evidence of learning based on infants' looking patterns. During the first and second windows of analysis, infants looked proportionally more toward the target object than the distractor object. The onset contingent plots also showed evidence of learning. During the first window of analysis, infants shifted off the distractor object and onto the target object more often. If infants were already fixated on the target object during label onset, they tended to stay on it. In the second window of analysis, the same pattern emerged: infants shifted from the distractor object to the target object, and once fixated on the target object, they tended to stay there. Additionally, infants’ mean accuracy during the second window of analysis was significantly above the chance level. This pattern suggests that infants learned the novel labels in the varied condition when they created the varied views through their object manipulation.
Between the two experiments, a marginally significant trend indicated that mean accuracy tended to be higher in the varied conditions than in the fixed conditions. There was no interaction between the experiments and conditions or between the experiments and the analysis window. Still, there was a trend that infants had higher mean accuracies when objects had multiple varied views than when they had fixed, consistent views. The varied view condition in the second experiment is the only novel word condition with a critical window above the chance level. Overall, the findings of this study suggest that manipulating objects in a way that creates varied views while hearing the novel label was the best condition for word learning.

**Individual differences and learning**

In the first experiment, the individual infant’s engagement in task dynamics also did not impact the rates of learning. The proportion of looking at the scene during the learning trials was similar between the familiar and novel words in both the fixed and varied conditions. Furthermore, the proportion of looking at the scene was not correlated with mean accuracies during the test trial. This has two implications. First, infants engaged with the task similarly regardless of how the object was presented. Second, the overall amount of looking and the experience of the varied views did not predict learning.

In the second experiment, rates of object manipulation were low, likely due to the artificial way the object was affixed to the apparatus, with infants' hands on the object for only about 30% of the trial. Additionally, infants manipulated the objects on average for the same proportion of time between the two conditions. Despite these similarities, learning was significantly different in the two conditions, with infants showing a higher mean accuracy in the varied condition compared to the fixed condition. This suggests that the multiple, varied views...
that infants created through manual exploration enhanced learning above and beyond the mere amount of object manipulation.

Surprisingly, there was no significant relationship between the proportion of time infants spent manipulating objects during learning trials and their mean accuracy. Infants were largely similar in the amount of time they spent manipulating objects, which suggests that the amount of object manipulation does not necessarily affect learning. However, there was a significant positive correlation between the proportion of time infants spent attending to the scene in the varied condition and their mean accuracy. This shows that the individual’s propensity to look at the views they created impacted their accuracy. Individual participant factors did not significantly impact mean accuracy, except for age, which negatively correlated with mean accuracy in the fixed condition. In summary, for the second experiment, the proportion of time spent manipulating objects did not impact learning. Still, the proportion of time spent attending to the scene was positively correlated with mean accuracy in the varied condition. This finding suggests that their accuracy was higher when infants observed more varied views that they created through their object manipulation.

It was interesting that the vocabulary scores did not predict learning in either of the experiments. There was a wide range in the variability of the vocabulary scores, indicating that the infants brought varying levels of word-learning expertise to the lab. Despite the variability in vocabulary scores, infants demonstrated similar patterns of learning. In the first experiment, most infants (five out of eight) had higher mean accuracies in the varied condition than the fixed. In the second experiment, the pattern was stronger, with six out of eight infants showing higher mean accuracies in the varied condition.
The role of varied object views in word learning

Few studies have examined the role of object manipulation and in-the-moment word learning. However, Slone et al. (2019) found that infants who manipulated objects in a way that produced varied views had larger vocabulary sizes six months later. Therefore, creating varied object views may be related to word learning. The present study is the first to tease apart the relative impact of manual exploration on word learning. It showed that infants have higher mean accuracies for target objects when they manually explore an object in a way that provides multiple, varied object views.

Understanding why varied object views may impact novel word learning is important. There is an argument that it could be driven by perceptual salience—objects that move in a way that creates varied views are more visually interesting. Therefore infants will attend to those views more. Pruden and colleagues (2006) found that 10-month-olds could learn novel word labels (one-syllable labels) when the novel object was perceptually salient and not boring. For example, a multicolored party shaker was a perceptually salient, therefore interesting object, whereas a white cabinet latch was considered boring. Furthermore, Pruden et al. found that infants spent more time looking at the perceptually interesting object during the study's training phase than the boring object. In the present study, the varied movement could create more salient moments of attention, thus, higher accuracy for the varied objects in the experiment with manipulation.

However, if saliency was the main driving factor for these results, we would expect to see a greater proportion of time spent looking at the scene in the varied condition than in the fixed condition, as was seen in Pruden et al., (2006). However, there was not a difference in the overall proportion of looking to the scene between conditions or across experiments. While there was an
effect that the amount of time spent looking at the scene during the learning trials for the varied condition in the experiment with manipulation had higher mean accuracies during the test trials, there was no difference in the proportion of time spent attending the scene between the fixed and varied conditions. Therefore, it is more likely that the active experience of varied object views impacted learning, not just the relative salience between the conditions during the learning trials.

The visual experience between the two experiments was also different. In Experiment 1, the experimenter had their hand on the object for the trials and rotated it at a fixed, consistent pace. When infants could manually explore the object in Experiment 2, they tended to do so in a much more sporadic fashion. Their hands were only on the objects about 30% of the trial, and they tended to manipulate the object in a way that would allow it to continue to rotate even if their hands were not on it. This difference in the visual experience between the two experiments is important to consider when interpreting the results. The more consistent visual experience in Experiment 1 may have limited the infants’ ability to explore and learn about the object from different angles, leading to the weaker learning response seen in this experiment. Whereas the more dynamic visual experience in Experiment 2 may have allowed for a richer and more comprehensive understanding of the object.

In contrast to infant created movement, the consistent rotation present in Experiment 1 could have been a better learning environment for the infants. However, there was still inconsistent evidence for learning potentially due to the lack of creating the views independently. The lack of clear evidence for learning with fixed object views is line with the findings of Harman et al. (1999). In this study, adults had better object recognition if they created object rotations compared to passively observing them. This interpretation of the results is further
supported by the fact that there was no significant difference in the proportion of time infants spent looking at the scene for each condition between the two experiments.

Object recognition may be a driving factor that explains why infants had higher mean accuracies for the varied objects than the fixed in Experiment 2. Orr (2020) found that object exploration mediated gesture and speech production. Furthermore, Orr suggested that the ability to recognize objects is the key factor driving this relationship. Accordingly, the connection between manipulating objects and object naming could be primarily driven by the mechanism of object recognition, which involves recalling information about the object.

Prior research has shown that object recognition is connected to the creation of planar object views (James et al., 2014). In the present study, the angle of the object views was consistent, using planar views. This means they were presented perpendicular to their primary axis (i.e., not along a diagonal rotation). The different views were also planar when the varied view objects were rotated. In the second experiment, the difference between learning in the fixed and varied conditions does not just come down to the type of view, for example, planar or disorganized, but the variety of views seen. Thus, experiencing a variety of planar views may improve word learning. Returning to Horst and Samuelson's (2008) five tasks infants need to do to learn a word, the second to last step before encoding the word form is to encode something about the referent. Therefore, perhaps in this study, varied views experienced through object manipulation created opportune moments to encode different features of the object, thus providing a better learning environment.

**Implications**

This study is the first to link differing object views created through object manipulation to in-the-moment novel word learning. Thus far, much of the work has looked at correlations
between object views or manipulation and later vocabulary outcomes. However, this study used a controlled experimental setup to isolate the relative impact of object manipulation and varied object views. These results have implications for an expanded theoretical approach to novel word learning. As discussed earlier, Samuelson (2021) called for an updated theoretical perspective on how infants learn words, pointing out that many current theoretical models are inadequate and lack an understanding of how complex systems interact to produce learning. This study had a clear interplay between varied views, object manipulation, and word learning.

The present study's findings lend further support to the idea that cognition is inherently embodied, a key tenant of dynamic systems theory. This theoretical approach emphasizes the bidirectional relationship between the brain, body, and environment. The results of the present study suggest that infants’ own object manipulation may play a key role in creating an optimal learning environment. When infants were able to create varied views of the object, it resulted in more effective word learning. These findings align with earlier research highlighting the importance of the interplay between object manipulation and simultaneous exposure to novel labels for effective word learning (DiMercurio, 2019; Schroer & Yu, 2022; Yu & Smith, 2012).

Notably, other theoretical approaches would not fully account for the present findings. The intersensory redundancy hypothesis would posit that redundant information across senses should heighten learning by drawing attention towards the event (Gogate & Bahrick, 2001). If redundancy were the sole explanation for these findings, then we would not expect to see a difference between the fixed and varied conditions in the experiment with infant directed object manipulation. In this experiment, they received dual sensory information while touching the object and hearing the label. Instead, there was a condition difference with the varied views relating to higher mean accuracies. Therefore, it was the dynamic interweaving of hearing the
label, manipulating the object, and seeing the varied views that impacted novel word learning. Moreover, these findings highlight how the body and environment can work together to facilitate learning a key component to dynamic systems theory.

This work also has potential implications for applied work. Iverson and Braddock (2011) noted that children with specific language impairments also have gross and fine motor development delays. While this could be related to global delays in development, the deficits in the motor system could lead to delays in language development. Previously discussed work, and this study, have demonstrated that object manipulation plays a definitive role in word learning. Future work could look at how interventions related to motor skills impact later language development.

Limitations and future directions

While the significant results and trends that emerged in this study are noteworthy, the sample size is a limitation. The sample was only half of what was needed to reach power; therefore, the study is underpowered. It is worth noting that even with a small sample size, significant effects, and trends still emerged in the hypothesized direction. This may be due to the controlled experimental design used in the study. Nonetheless, it is important to replicate these findings with larger sample sizes to ensure that the results are reliable. In addition, a larger sample size would allow for a greater exploration into variability in learning and potential individual differences.

It is worth noting that both conditions in Experiment 1 showed similar delays in visually shifting to the target object, which suggests that the task was challenging for the infants, and it took them longer to fixate on the target object. The delays make the results of this study difficult to interpret and present a limitation. Yet, there was not a delay for the familiar words. Therefore,
it is likely that the task for the novel words was challenging, and these conditions presented weak evidence for learning. Alternatively, infants could display an alternative pattern to looking outside what is typically seen in a canonical looking while listening. Although, infants did demonstrate the ability to select familiar words accurately when presented in a visual-only setting. Therefore, it is possible that the novel words presented a more difficult learning environment for infants.

While there was no difference in the proportion of time infants spent looking at the scene between conditions and across experiments, the inherent difference in how the objects were rotated could be a limitation. In Experiment 1, the objects' rotations were fairly consistent, but in Experiment 2, infants created sporadic rotations. Because of this difference, it is difficult to interpret the relative impact on learning. However, the sporadic rotations in Experiment 2 had better rates of learning in the varied condition.

This study was also not naturalistic. The experimental setup used in this study was highly controlled, which allowed us to understand the relative impact of manipulation and object views. However, this also means that the setup does not entirely represent how infants experience objects daily. In particular, affixing the objects to a rod limited how infants could manipulate them. For example, when infants manually explore objects in their daily lives, they can move them in various ways, often along multiple axes and in different directions. In contrast, in Experiment 1 of this study, the experimenter had their hand on the object for the duration of the trials and rotated it at a fixed, consistent pace. While this allowed for a consistent presentation of the object, it does not reflect how infants would typically explore objects in their environment.

Similarly, in Experiment 2, while infants could manually explore the objects, they tended to do so in a much more sporadic fashion than the controlled rotations in Experiment 1. The
infants' hands were only on the objects about 30% of the trial, and they tended to manipulate the object in a way that would allow it to continue to rotate even if their hands were not on it. While this allowed for manual exploration, it was still constrained by the experimental setup. In reality, if infants actively explore an object, their hands would typically be on the object for most of the bout.

It should also be noted that while it was interesting that there are condition differences in the second experiment despite the use of planar views, this may be a limitation. This study used planar views, which are fairly similar to the views infants typically create at this age. However, it is unclear if the angle of the varied and fixed views would have impacted learning. If the views were varied but disorganized, for example, on a diagonal axis, there might not have been a relationship between the views and word learning.

Future work should examine the role of varied object views from object manipulation during word learning in a more naturalistic setting. While Schroer and Yu's (2022) work was naturalistic, it leaves out a key part of understanding why multimodal attention impacted learning novel object labels. The authors suggest that it could be due to creating richer visual data while manipulating the objects. Still, the type of visual data infants created in this study is unclear. In addition, future work should examine how infants move and rotate objects during a label learning task in which they have full freedom of movement. This would shed light on how varied object views relate to word learning in day-to-day play.

Additionally, future work should examine how varied and fixed object views impact object recognition in younger infants regarding novel word learning. Examining this area could help explain why there is a link between object exploration and novel word learning. In this study, objects across conditions were never tested against each other, so there was no way, post-
hoc, to see if differences in the experienced views impacted object recognition. Future studies with infants of the same age could develop an additional component in a word-learning task to measure object recognition.

Finally, while it is notable that the field is shifting to understand how motor and language skills develop together, much of the research is still correlational. Often the focus is on studying factors, such as motor skills or environmental exploration, concerning broader outcomes, typically vocabulary size, as reported by the MCDI. This work has led to exceptional findings detailing developmental cascades and the complex nature of development. However, this approach overlooks the processes underlying development and solely focuses on the outcomes. Therefore, while balancing correlational work that allows for a naturalistic look at development, more work should focus on the underlying processes involved in the relationship between motor skills and language development.

Conclusion

The two experiments investigated different methods of learning novel words in infants. In Experiment 1, infants passively viewed the objects and heard the novel word labels in fixed and varied view conditions. The results showed a slightly stronger learning response in the varied view condition, with higher shifts from the distractor to the target and lower shifts away from the target, indicating that the multiple, varied views of the novel object presented to the infants may have boosted their learning rates in a challenging environment that allowed for only visual exploration. However, this was only a trend and not statistically significant. In addition, the pattern of looking across the conditions showed a delayed shift to the target object. Similar engagement in the task dynamics between familiar and novel words suggests that the task was challenging for the infants. It took them longer to focus on the target object.
In Experiment 2, infants manually explored objects while hearing novel word labels in fixed and varied view conditions. The results showed that despite low rates of object manipulation, infants showed significantly higher mean accuracy in the varied condition, indicating that the multiple, varied views that infants create through their manual exploration can enhance word learning. Importantly, attending to the varied objects more was related to higher mean accuracies. Experiment 1 suggested that varied object views could potentially boost learning rates in challenging visual-only environments. In contrast, Experiment 2 showed that infants' manual exploration and creation of varied object views could significantly enhance word learning.

These two experiments provided novel insight into the role of varied object views created through object manipulation in word learning. The findings suggest that allowing infants to create varied views through self-directed object manipulation can significantly enhance their accuracy in a novel word-learning task. These results underscore the importance of infants' active role in their learning. Additionally, these findings highlight the potential benefits of providing infants with opportunities to actively explore their environment to support their language development. Overall, this study contributes to understanding the complex system underlying infants' word learning while manipulating objects.
References


https://doi.org/10.1123/mcj.6.1.52


https://doi.org/10.1037/0012-1649.29.5.832


https://doi.org/10.1186/s11689-018-9231-3


Appendix

Table 1. Sociodemographic information for participants.

<table>
<thead>
<tr>
<th></th>
<th>Experiment 1</th>
<th></th>
<th>Experiment 2</th>
<th></th>
<th>Full Sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>5</td>
<td>62.5</td>
<td>4</td>
<td>50</td>
<td>9</td>
<td>56.25</td>
</tr>
<tr>
<td>Male</td>
<td>3</td>
<td>37.5</td>
<td>4</td>
<td>50</td>
<td>7</td>
<td>43.75</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>8</td>
<td>100</td>
<td>7</td>
<td>87.5</td>
<td>15</td>
<td>93.75</td>
</tr>
<tr>
<td>Mixed</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>12.5</td>
<td>1</td>
<td>6.25</td>
</tr>
<tr>
<td>Parental Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>0</td>
<td></td>
<td>1</td>
<td>12.5</td>
<td>1</td>
<td>6.25</td>
</tr>
<tr>
<td>Associate/Bachelors</td>
<td>4</td>
<td>50</td>
<td>1</td>
<td>12.5</td>
<td>5</td>
<td>31.25</td>
</tr>
<tr>
<td>Masters/Professional/Doctorate</td>
<td>4</td>
<td>50</td>
<td>6</td>
<td>75</td>
<td>10</td>
<td>62.5</td>
</tr>
<tr>
<td>Parent 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Associate/Bachelors</td>
<td>3</td>
<td>37.5</td>
<td>6</td>
<td>75</td>
<td>9</td>
<td>56.25</td>
</tr>
<tr>
<td>Masters/Professional/Doctorate</td>
<td>5</td>
<td>62.5</td>
<td>2</td>
<td>25</td>
<td>7</td>
<td>43.75</td>
</tr>
</tbody>
</table>
Table 2. Example of Trial Ordering for Learning Trials

<table>
<thead>
<tr>
<th>Block</th>
<th>Trial</th>
<th>Condition</th>
<th>Stimulus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Familiar</td>
<td>Duck</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Familiar</td>
<td>Shoe</td>
</tr>
<tr>
<td>Learn</td>
<td>3</td>
<td>Varied</td>
<td>Modi</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>Fixed</td>
<td>Teema</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>Varied</td>
<td>Nilla</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>Fixed</td>
<td>Dobu</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>Familiar</td>
<td>Duck</td>
</tr>
<tr>
<td>Learn</td>
<td>8</td>
<td>Varied</td>
<td>Modi</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>Fixed</td>
<td>Teema</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>Varied</td>
<td>Nilla</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>Fixed</td>
<td>Dobu</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>Familiar</td>
<td>Shoe</td>
</tr>
<tr>
<td>Learn</td>
<td>13</td>
<td>Varied</td>
<td>Nilla</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>Fixed</td>
<td>Dobu</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>Fixed</td>
<td>Teema</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>Varied</td>
<td>Modi</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>Familiar</td>
<td>Shoe</td>
</tr>
<tr>
<td>Learn</td>
<td>18</td>
<td>Familiar</td>
<td>Duck</td>
</tr>
<tr>
<td>4</td>
<td>19</td>
<td>Varied</td>
<td>Nilla</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>Fixed</td>
<td>Teema</td>
</tr>
<tr>
<td>21</td>
<td>21</td>
<td>Fixed</td>
<td>Dobu</td>
</tr>
<tr>
<td>22</td>
<td>22</td>
<td>Varied</td>
<td>Modi</td>
</tr>
</tbody>
</table>
Table 3. Example of Test Trial Ordering

<table>
<thead>
<tr>
<th>Block</th>
<th>Trial</th>
<th>Target</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>23</td>
<td>Shoe</td>
<td>Shoe - Duck</td>
</tr>
<tr>
<td>Test 1</td>
<td>24</td>
<td>Duck</td>
<td>Shoe - Duck</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>Teema</td>
<td>Dobu - Teema</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>Nilla</td>
<td>Modi - Nilla</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>Modi</td>
<td>Modi - Nilla</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>Dobu</td>
<td>Dobu - Teema</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>Duck</td>
<td>Duck - Shoe</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>Modi</td>
<td>Modi - Nilla</td>
</tr>
<tr>
<td>Test 2</td>
<td>31</td>
<td>Dobu</td>
<td>Teema - Dobu</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>Nilla</td>
<td>Nilla - Modi</td>
</tr>
<tr>
<td></td>
<td>33</td>
<td>Teema</td>
<td>Teema - Dobu</td>
</tr>
<tr>
<td></td>
<td>34</td>
<td>Duck</td>
<td>Duck - Shoe</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>Shoe</td>
<td>Duck - Shoe</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>Dobu</td>
<td>Dobu - Teema</td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>Nilla</td>
<td>Modi - Nilla</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>Teema</td>
<td>Teema - Dobu</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>Modi</td>
<td>Nilla - Modi</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>Shoe</td>
<td>Shoe - Duck</td>
</tr>
<tr>
<td></td>
<td>41</td>
<td>Dobu</td>
<td>Dobu - Teema</td>
</tr>
<tr>
<td>Test 4</td>
<td>42</td>
<td>Teema</td>
<td>Teema - Dobu</td>
</tr>
<tr>
<td></td>
<td>43</td>
<td>Modi</td>
<td>Nilla - Modi</td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>Nilla</td>
<td>Modi - Nilla</td>
</tr>
</tbody>
</table>

Note. The Bold items in the Order column represent the target object and the side it should be presented on.
Figure 1. Experimental Set-up. (A) Set up for the learning phase trials. For experiment 1, there was a plexi-glass sheet in front of the object. (B). Test trial set up. For both experiments, there was a plexi-glass sheet in front of the objects.
Figure 2. Stimuli for Familiar Words.
Figure 3. Stimuli for Novel Words. The rod ran through the middle of the object, horizontally.

(A) Novel stimuli in the position for fixed views. (B) Novel stimuli for varied views.
Figure 4. Timeline of Learning Trials. This shows the timeline of the auditory stimulus for learning trials. It was consistent across experiments. The same type of string was used for novel and familiar words. The dashed lines and times represent the label onset.
Figure 5. Timeline of Test Trials. This shows the timeline of the auditory stimulus for test trials. It was consistent across experiments. The same type of string was used for novel and familiar words. There was silence for the first 1500ms. The dashed lines at 2000 and 4500ms represent each label onset. The grey box represents the window of analysis.
Figure 6. Looking During Learning Trials (Experiment 1). The proportion of looking to the target object during the learning trials across the familiar and novel word conditions. The error bars show standard error.
Figure 7. Familiar Words Mean Accuracy During Windows of Analysis (Experiment 1). Mean accuracy by the window of analysis. The dashed line represents chance level. The error bars show the standard deviation.
Figure 8. Familiar Words Time Course (Experiment 1). The dashed horizontal line shows chance level while the dashed vertical lines show the label onsets. The time started with the first label onset while the second label onset started at 2500ms. The first 2000ms before label onset represent the baseline window. The two shaded areas represent the two windows of analysis. The solid, green line is the average proportion looking to the target. The shaded area around the average shows the standard error.
Figure 9. Familiar Words Onset Contingent Plot (Experiment 1). The dashed horizontal line shows chance level. The two shaded areas represent the two windows of analysis. The solid, red or blue line is the average proportion shifting. The shaded area around the average shows the standard error.
Figure 10. Novel Words Mean Accuracy During Windows of Analysis (Experiment 1). Mean accuracy in each window of analysis. The dashed line represents chance level. The error bars show standard error.
Figure 11. Novel Words Mean Accuracy for Conditions (Experiment 1). Mean accuracy in each condition. The dashed line represents chance level. The error bars show standard error.
Figure 12. Novel Words Mean Accuracy for Condition by Window (Experiment 1). Mean accuracy for between conditions and window of analysis. The dashed line represents chance level. The error bars show standard error.
Figure 13. Novel Words Time Course (Experiment 1). The dashed horizontal line shows chance level while the dashed vertical lines show the label onsets. The time started with the first label onset while the second label onset started at 2500ms. The first 2000ms before label onset represent the baseline window. The two shaded areas represent the two windows of analysis. The solid, red or blue line is the average proportion looking to the target. The shaded area around the average shows the standard error.
Figure 14. Novel Words Onset Contingent Plot (Experiment 1). The dashed horizontal line shows chance level. The two shaded areas represent the two windows of analysis. The solid, red or blue line is the average proportion shifting. The shaded area around the average shows the standard error.
Figure 15. Novel Words Variability (Experiment 1). The dashed line represents chance level. Each line represents a participant and their individual mean accuracies.
Figure 16. Familiar Words Correlations (Experiment 1). The dashed line represents the trend for the correlations.
Figure 17. Fixed Condition Correlations (Experiment 1). The dashed line represents the trend for the correlations.
Figure 18. Varied Condition Correlations (Experiment 1). The dashed line represents the trend for the correlations.
Figure 19. Looking Correlations with Mean Accuracy (Experiment 1). The dashed line represents the trend for the correlations.
Figure 20. Proportion Looking During Learning Trials (Experiment 2). The proportion of looking to the target object during the learning trials across the familiar and novel word conditions. The error bars show standard error.
Figure 21. Proportion Manipulation During Learning Trials (Experiment 2). This graph shows the proportion of time infants spent manipulating the objects. The error bars represent standard error.
Figure 22. Familiar Words Mean Accuracy During Windows of Analysis (Experiment 2). Mean accuracy in each window of analysis. The dashed line represents chance level. The error bars show the standard deviation.
Figure 23. Familiar Words Time Course (Experiment 2). The dashed horizontal line shows chance level while the dashed vertical lines shows the label onsets. The time started with the first label onset while the second label onset started at 2500ms. The first 2000ms before label onset represent the baseline window. The two shaded areas represent the two windows of analysis. The solid, green line is the average proportion looking to the target. The shaded area around the average shows the standard error.
Figure 24. Familiar Words Onset Contingent Plot (Experiment 2). The dashed horizontal line shows chance level. The shaded area represents the window of analysis. The solid, red or blue line is the average proportion shifting. The shaded area around the average shows the standard error.
Figure 25. Novel Words Mean Accuracy During Windows of Analysis (Experiment 2). Mean accuracy in each window of analysis. The dashed line represents chance level. The error bars show standard error.
Figure 26. Novel Words Mean Accuracy for Conditions (Experiment 2). Main effect of condition on mean accuracy. The dashed line represents chance level. The error bars show standard error. The asterisk shows that there was a significant difference between these conditions.
Figure 27. Novel Words Mean Accuracy for Condition by Window (Experiment 2). Mean accuracy by condition and window of analysis. Varied in the second window is significantly above chance level. The dashed line represents chance level. The error bars show standard error.
Figure 28. Novel Words Time Course (Experiment 2). The dashed horizontal line shows chance level while the dashed vertical lines shows the label onsets. The time started with the first label onset while the second label onset started at 2500ms. The first 2000ms before label onset represent the baseline window. The two shaded areas represent the two windows of analysis. The solid, red or blue line is the average proportion looking to the target. The shaded area around the average shows the standard error.
Figure 29. Novel Words Onset Contingent Plot (Experiment 2). The dashed horizontal line shows chance level. The shaded area represents the window of analysis. The solid, red or blue line is the average proportion shifting. The shaded area around the average shows the standard error.
Figure 30. Novel Words Variability (Experiment 2). The dashed line represents chance level. Each line represents a participant and their individual mean accuracies.
Figure 31. Familiar Words Correlations (Experiment 2). The dashed line represents the trend for the correlations.
Figure 32. Fixed Condition Correlations (Experiment 2). The dashed line represents the trend for the correlations. $p < .05$ notes that this plot shows a significant correlation.
Figure 33. Varied Condition Correlations (Experiment 2). The dashed line represents the trend for the correlations.
"Figure 34. Looking Correlations with Mean Accuracy (Experiment 2). The dashed line represents the trend for the correlations. $p < .05$ notes that this plot shows a significant correlation."
Figure 35. Manipulation Correlations with Mean Accuracy (Experiment 2). The dashed line represents the trend for the correlations.
Figure 36. Proportion Looking for Experiments (Between-Subjects). Proportion of looking during the learning trials. The error bars show standard error.
Figure 37. Mean accuracy for Experiments (Between-Subjects). Mean accuracy during the test trials between experiments. The dashed line represents chance level. The error bars show standard error.
Figure 38. Mean accuracy for Window of Analysis (Between-Subjects). Mean accuracy between the windows of analysis. The dashed line represents chance level. The error bars show standard error.
Figure 39. Mean accuracy Condition (Between-Subjects). Mean accuracy by conditions across experiments. The dashed line represents chance level. The error bars show standard error.
Figure 40. Mean accuracy for Condition by Experiment (Between-Subjects). Mean accuracy between conditions and experiments. The dashed line represents chance level. The error bars show standard error.
Vita

Abigail DiMercurio was born in Atlanta, Georgia to Debbie and Sam DiMercurio. She is a sister to Nicole DiMercurio. She attended Riverside Elementary and continued to Lanier Middle and North Gwinnett Highschool in Suwanee, Georgia. After graduation, she attended the University of Georgia, where she was introduced to developmental psychology and experimental research. Abigail started developmental research with Dr. Janet Frick. She obtained a Bachelor of Science degree with the University of Georgia at Greensboro in May 2016 in Psychology with a minor in Child and Family Studies. She accepted a graduate research assistantship at the University of Tennessee, Knoxville in the experimental psychology program in the Infant Perception-Action Laboratory, directed by Dr. Daniela Corbetta. Abigail received a Master of Arts degree in Psychology at the University of Tennessee Knoxville in May 2019 and continued at the University of Tennessee, Knoxville to pursue her Doctor of Philosophy degree in Experimental Psychology.