

The Influence of Task Dynamics on Object-Label Mapping: Synchrony Between Perception and Action

A Thesis Presented for the
Master of Arts
Degree
The University of Tennessee, Knoxville

Abigail Julian DiMercurio
May 2019

ACKNOWLEDGEMENTS

First, I want to thank my advisor, Daniela Corbetta, for her support, guidance, and expertise throughout this project. I also want to thank my committee, Jessica Hay and Shannon Ross-Sheehy, for their input and feedback on this project. Additionally, I want to thank my fellow labmates, John Connell, Matthew Clark and Rebecca Wiener in their assistance in data collection. I would not have data without their help. I also would like to thank two undergraduate research assistants for doing the reliability coding for my data, Louma Hijer and Emily Medford. Lastly, I need to thank my parents, Sam and Debbie DiMercurio, as well as my sister, Nicole DiMercurio, for their constant love and support.

ABSTRACT

Often areas of infant development such as action and cognition are studied separately, despite fine interactions between the two developmental systems. Specifically, there is a relationship between motor skills and language development seen with a robust, positive correlation between more advanced locomotor skills and larger vocabulary sizes. Additionally, when infants create a large number of different object views through manual object exploration, they also demonstrate larger vocabulary sizes later in development. The present study proposes a model of examining language skills, specifically object-label mapping using Dynamic Systems Theory with multileveled interactions between attention, infant manual exploration, and novel label timing feeding into better mapping of novel labels onto objects. The present study seeks to address how dynamics of a task involving object manipulation impact subsequent object-label mapping. Infants, aged 18-21 months, proceeded through three within-subject conditions varied in their degree of manipulation and label timing. Two groups of infants were pulled from the data based on their performance in a condition with no infant manipulation. The results demonstrate that infants who struggled with mapping novel labels onto the object during the condition with no manipulation readily learned when there was synchrony between looking at the object while manipulating and hearing the novel label. The current findings call for a more in-depth look into foundational learning processes during early language acquisition in order to explain the variability seen in language skills.

TABLE OF CONTENTS

CHAPTER ONE: INTRODUCTION	1
The Dynamic Relationship Between Action and Cognition	1
CHAPTER TWO: LITERATURE REVIEW	5
Language Development	5
Variability in Language Development	8
Motor Development and Environmental Exploration	9
Links Between Motor and Language Development	11
Object Exploration and Object-Label Mapping	14
Components of Task Performance Success	17
A Dynamic Systems Approach to Language	20
Purpose and Hypotheses	22
CHAPTER THREE: METHOD	25
Participants	25
Stimuli and Apparatus	25
Questionnaire and Vocabulary Assessment	26
Procedure	27
Look-Only Condition	29
Out-of-Sync Condition	29
In-Sync Condition	29
Test Trials	30
Coding and Analyses	30
CHAPTER FOUR: RESULTS	34
Presentation Trial Performance	34
Proportion of Looking During Presentation Trials	35

Looking While Hearing During Presentation Trials	35
Manipulation During Presentation Trials.....	36
Manipulation While Looking During Presentation Trials	36
Performance During Test Trials.....	37
Performance Between the Learning Groups	37
Study Dynamics and Performance.....	38
Age and Performance.....	38
MCDI and Performance	39
Walking Onset and Performance	39
Talking Onset and Performance.....	40
Looking During Presentation and Performance.....	40
Looking and Hearing During Presentation And Performance	40
Manipulation During Presentation and Performance.....	41
CHAPTER FIVE: DISCUSSION.....	42
A Dynamic View of Learning.....	42
The Role of Holding while Hearing.....	43
Environmental and Research Implications	46
Limitations and Future Directions	47
Conclusion	49
REFERENCES	51
APPENDICES	59
VITA.....	79

LIST OF TABLES

Table 1. Mean estimates between groups and condition for presentation looking	72
Table 2. Mean estimates for looking while hearing.....	73
Table 3. Mean estimates between groups and condition for looking while hearing.....	74
Table 4. Mean estimates for manipulation presentation trials	75
Table 5. Mean estimates for looking during manipulation during presentation trials	76
Table 6. Mean estimates between learning groups and condition	77
Table 7. Pairwise comparisons between the performance for group and condition	78

LIST OF FIGURES

Figure 1. Dynamic Systems Theory model for object-label mapping	60
Figure 2. Experimental setup	61
Figure 3. Novel stimuli.	62
Figure 4. Timeline of auditory string.	63
Figure 5. Description of the differences between the conditions.....	64
Figure 6. Looking preference during the baseline window of the test trial	65
Figure 7. Main effect of looking while hearing	66
Figure 8. Interaction between condition and group for looking while hearing	67
Figure 9. Main effect of manipulation	68
Figure 10. Main effect of looking during manipulation	69
Figure 11. Interaction between condition and group for test trials	70
Figure 12. Individual differences for test trial results.....	71

CHAPTER ONE: INTRODUCTION

The Dynamic Relationship Between Action and Cognition

During the first months of life, infants experience rapid developmental changes in the motor, language, social, and cognitive domains. Typically, these richly studied domains of infant development have been studied in isolation, with minimal cross over between subfields. Studying infant development in silos presents an issue when it comes to studying different processes of ontogeny, as these areas co-develop and continually interact. One classical and integrative theory, proposed by Jean Piaget (1970), highlights the role of changes in sensorimotor experiences and how the changes and interactions impact cognition. This theory proposed that during the first year of life, infants learn through sensorimotor experiences, their actions within their environment. Therefore, cognition is embodied, meaning that our knowledge of the world arises from the tight coupling of our body, which has interacting perceptual and motor systems, nested in our environment (Thelen, Schöner, Scheier, & Smith, 2001).

The embodied cognition perspective posits that action can impact learning across the lifespan, not just in infancy. Taking an embodied perspective highlights the importance of studying different domain of infant development together, especially regarding how infants' sensory interactions relate to their subsequent understanding of the environment. When viewing development in the fashion of isolated domains, we lose the finer interactions between multiple systems. We can expand the overall understanding of different facets of infant development by studying infants' interactions within their environment while adopting a dynamic systems view of development (Thelen and Smith, 1998).

The relationship between motor development and cognition provides an excellent example of how different systems intersect. For example, there is an established relationship between age of onset of motor skills and subsequent language development (Walle & Campos, 2014; He, Walle, & Campos, 2015; Oudgenoeg-Paz, Leseman, & Volman, 2015; Libertus & Violi, 2016). The mediator between the dynamic relationship between motor and language skills could be environmental exploration. For instance, as infants gain new motor skills, they interact with their environment at a higher rate (Thurman & Corbetta, 2017). A critical experience that occurs with increases in environmental exploration is an increase in sensory information obtained from the environment. Therefore, as infants experience greater environmental interactions, they are gaining more knowledge, thus developing their language skills.

An infant's environment can provide a plethora of information through the occurrence of continually reshaping perception-action (Gibson, 1958). That is, infants gain sensory information by interacting with their environment, and in turn, sensory feedback allows for learning object properties. The series of tasks described here is continually looping so that infants are exploring their environment, interacting with it, and learning from it. The loop of perception and action can impact how infants explore their environment (Franchak, van der Zalm, & Adolph, 2010). For example, Franchak and colleagues (2010) show that the sensory information infants gain from each interaction can shift how likely they are to interact with that object in future bouts of play based on properties they learned during exploration, such as weight or texture. The perceptual cycle described above is what infants are experiencing at increasing rates as they gain motor skills to navigate their environment, therefore interact with more objects. With this understanding, infants are also increasing their knowledge about the environment, potentially including object labels, when interacting with their environment.

Here I have provided a preliminary overview of how action and cognition dynamically interact with the aim to bring this intersection into the specific domain of language development. Viewing the development of language through the lens of dynamic systems theory could have significant benefits to the understanding of the individual differences (Fenson et al., 1994; Walle & Campos, 2014) that are often seen within studies examining aspects of language development, such as vocabulary growth. Every single action in the environment is related to a multitude of causal actions, as well as new perceptual associations (Thelen, 1989). While this is infrequently studied in infancy, several computational models have demonstrated a dynamic relationship between language and embodiment. Little and Sommer (2011) developed a computational model of development demonstrating a relationship between the environment and learning when it comes to object-label associations. The model demonstrated greater success in novel object-label associations if the model coupled their actions within the object-label mapping environment, therefore acting as an embodied agent (Little & Sommer, 2011). Additionally, robots show similar benefits for object-label mapping when they act within their environment in an embodied way (Salvi, Montesano, Bernardino, & Santos-Victor, 2012). Both models demonstrated higher rates of learning about their environments and object-label associations when allowed to have embodied moments of learning.

Within the infant literature, however, few studies, have explored the role of perception and action in object-label learning tasks. What remains unanswered is how interactions with objects in the environment impact subsequent, novel object-label mapping. The present study seeks to fill this gap in the literature by investigating how infant's interactions with novel objects and the timing of novel word label production has an impact on object-label mapping. We are also assessing if individual differences in object-label mapping can explain the

variability we see with language learning. To begin to answer this question, I will first discuss the current fields of language and motor development, as it relates to environmental exploration. Then I will bring the two fields together with a discussion of the few studies that have linked perception, action, and language development before proposing the current study.

CHAPTER TWO: LITERATURE REVIEW

Language Development

There are many different theories of language acquisition that, for example, stress innate mechanisms or the importance of the social environment. These different views typically exclude the infant as the primary actor in the environment; however, understanding the different theories allows for a foundational look towards a more integrative approach to language development. An integrative theory should be moving towards highlighting infants as active agents in their environment, as well as the role of infants' perceptual associations with objects in their environment. A historical, but provocative theory on language development is that language is something that appears without formal instruction, unlike other cognitive skills such as mathematics and science, and therefore must be an innate mechanism—a language acquisition device (Chomsky, 1967). Although this perspective on language acquisition has been influential, the theory does not justify the individual differences present in language acquisition, which prompted a great surge of research into alternative explanations of language development.

Another well-cited perspective on language development involves the social relationships and social environments of infants. This theory is centered around the role of the caregiver in guiding the development of language through repetitive social contexts. The social theory of word learning derives from the issue called the poverty of the stimulus (Quine, 1964). Poverty of the stimulus highlights the complex, at times overwhelming environment infants are in. In almost any context, there is an overwhelming number of stimuli with ambiguous referents. The social theory of language acquisition allows for some interpretation into a complex issue of object-label mapping, classically demonstrated in the “gavagai” example (Quine, 1960). This

complex mapping issue involves an individual surrounded by those who speak an unfamiliar language trying to communicate about an ambiguous referent. The foreign language speaker would say “gavagai” in the presence of a stimulus, and the foreign language speakers point to a rabbit. However, they are also in the presence of hunting tools. The unfamiliar speaker recognizes that “gavagai” is a word, yet the individual cannot disambiguate the referent from the possible options. “Gavagai” could represent the rabbit, or it could be a command to hunt the rabbit, or it simply could be descriptive of the rabbit such as “elegant”.

The referentially ambiguous “gavagai” is similar to what infants experience during each day as they experience their native, yet novel language (Tomasello, 2010). The narrowing into words from a social partner is the social-pragmatic theory of language acquisition (Bruner, 1985). Under this theory, infants gradually learn novel words from associations within the environment of social input. An infant who is being carried out to the car and is told “let’s get in the car”, will soon recognize and associate, in similar scenarios, that what they are getting into is a car, not a stroller, nor a crib, due to the input from their social partner. However, this theory does not address differences in rates of language abilities for infants that share a similar language learning environment, therefore similar referential inputs (Walle & Campos, 2014). These differences in language abilities point to another explanation of how language develops. However, this is not to say that the social-pragmatic theory is not beneficial for language learners; it just is likely not a complete explanation for how infants learn object-label mappings.

Alternative theories of language acquisition center on a domain-general approach, such that perhaps there is not one central mechanism or theory for language acquisition but several, non-language specific cognitive processes that help the infant understand their environment (Samuelson & McMurray, 2017). This perspective moves the field away from viewing the onset

of language as happening through an innate, singular language acquisition device (Chomsky, 2014), or a domain-specific theory such as the social-pragmatic theory of language acquisition.

An example of domain-general learning is statistical learning which involves individuals picking up on regularities in a stream of information. With language specifically, infants can segment words in a speech stream due to the probabilistic regularities in the individual's natural language (Saffran, Aslin, & Newport, 1996). Statistical learning has been shown to promote novel object-label mapping by extracting regularities out of a speech stream (Fiser & Aslin, 2001; Pelucchi, Hay, & Saffran, 2009). Kirkham, Slemmer, & Johnson (2002) demonstrated that statistical learning is not specific to just language, but is a domain-general tool to extract regularities out of visual arrays as well. Samuelson and McMurray (2017) speculate that children may have the ability to learn words with such ease, not due to innate abilities, as seen with Chomsky's theory, but due to word learning as a process that is flexibly used across other cognitive areas, as seen with statistical learning.

A domain-general approach for language development, specifically in the realm of object-label mapping could be learning from a perception-action loop. A proposed perception-action loop of language learning could be highly relevant to the field of infant language acquisition; as infants physically interact with objects within their environment, they are receiving a high amount of sensory information. The subsequent increases in the processing of sensory information related to an increased number of perceptual associations (Thelen, 1989). The present study aims to understand if the increased perceptual associations with objects, gained through manual exploration, will have an impact on novel object-label learning.

The present state of infant language development research rarely involves the infant and the physical environment as a dynamically interacting system. Very few studies have considered

the role of the environment, beyond the social environment, in the study of language acquisition. However, those who have integrated action into language studies show interesting and compelling effects. Research on early language development with 6-month-olds show that infants demonstrated greater success in a visual selection task using familiar objects when the infants received more object labels that co-occurred with labeled object being physically present in the room (Bergelson & Aslin, 2017). While this specific study did not measure the infant's physical interaction with objects in the room, it shows that infants are learning words at a higher rate if the labeling occurred while infants were within sight of the object. These findings demonstrate the importance of the environment, and perhaps engagement with the environment in language development (see a discussion of Yu and Smith's work in the "Object Exploration and Object-Label Mapping" section for a more in-depth view of the intersection between object exploration and word learning).

Variability in Language Development

Examining language from a broader domain-general approach could begin to account for the differences we see at the level of the individual when learning language. Within the study of language development, there tends to be variability in different aspects of language such as vocabulary size, which can form a large gap between infants of the same age. For example, when testing the number of words infants understood (i.e., their receptive vocabulary), 16-month-olds can range from about 75 to nearly 350 words, according to parental report (Fenson et al., 1994). The wide-ranging vocabulary scores demonstrate a wide range of variability in how rapidly infants may obtain knowledge of words. As seen in Walle and Campos (2014), the individual's level of motor skill, such as if they are crawling or walking, can explain the variability in language outcomes, such that infants with more advanced motor skills also see higher vocabulary scores. Others propose that the difference is due to the quality of the

language input the infant receives (Peters, 1977). Another explanation nested in object exploration centers around the infant's likelihood of creating dynamic object-views, such that if they manipulate the object in a way that creates more individual views of the object, they have a larger vocabulary score later in development (Slone, Smith, & Yu, 2019).

Early work in individual differences cites that caregiver input (e.g., use of single nouns) typically matches the child's language production (Nelson, 1973). Furthermore, environmental differences, such as socioeconomic status can impact language development and vocabulary measures (Hoff, 2003). For example, Hart and Risely (1995) estimate that by the time toddlers are 3-years-old there is a 30 million word gap in the speech input the toddler receives if they are of a lower socioeconomic status. Examining aspects of language development in conjunction with other domains of development could help explain individual differences.

The previously described variability has looked at the impact of individual differences on the outcome of language development. However, it is essential to also understand individual differences in terms of the process of language learning, not just the outcome of vocabulary size. In order to understand development fully, it is essential to understand the process of development, and not just how factors impact the outcome of development. The need to understand the process highlights the importance of incorporating a dynamic-systems approach to understanding how interacting systems in the process of learning impact language development. Without understanding the dynamic interactions that may occur during the process of learning an object-label map, it will remain difficult to interpret the factors impacting individual variability in language outcomes.

Motor Development and Environmental Exploration

Gains in motor development are related to gains in environmental exploration (Thurman and Corbetta, 2017). Newborns are passive and cannot truly manipulate or explore their

environment, beyond the act of crying, which triggers action from their caregiver. At this stage, infants are just learning how their bodies and actions fit into their own space, which forms the foundation for the important motor milestone of reaching (DiMercurio, Connell, Clark, & Corbetta, 2018). Before the ability to reach, infants are only able to sparsely contact objects that are in their direct proximity, such as the mattress of their crib or a toy a caregiver has placed beside them. With the onset of reaching, infants are now able to make direct and deliberate reaches towards objects in their environment (Thelen et al., 1993). This gain in environmental exploration is crucial because it affords infants the chance to receive new sensory stimuli and begin to act on their surroundings. We can imagine an example with the perception-action loop; the ability to reach for and touch an object, triggered by an initial visual experience, can relate to a tangible grasp of the object and subsequent manipulation (Corbetta & Snapp-Childs, 2009). Furthermore, perception-action loops can be used to facilitate learning about object properties and how one's interaction with objects impacts the environment (Corbetta, DiMercurio, Wiener, Connell, & Clark, 2018b; Corbetta & Snapp-Childs, 2009). Therefore, reaching allows for further cycles of perception and action, in which infants can experience significant amounts of learning, in higher frequency than prior to the onset of reaching.

One of the next significant milestones to develop is the onset of crawling. Before this motor milestone, infants are only able to select objects in their proximity with a reach; however, infants cannot yet efficiently navigate towards the object they want. Infants are restricted to experiencing movement through passive means, such as a stroller or the arms of their caregiver. In these scenarios, infants have very little say in where they are going. Thurman and Corbetta (2017) related the onset of crawling to an increase in environmental exploration. The researchers followed a cohort of infants longitudinally across motor milestones. With the onset

of crawling, there was a significant increase in the distance they traveled in a playroom. Importantly, with the increase in environmental exploration, there was also an increase in the number of interactive bouts the infant had with objects in the environment (Thurman & Corbetta, 2017). The increased number of interactive bouts relates to an increase in sensory experiences; therefore, considering perception-action loops, the opportunity to learn from the environment increases as well.

Links Between Motor and Language Development

Previous research has demonstrated a relationship between motor and language development. Walle and Campos (2014) followed a cohort of infants across the transition of motor milestones, from crawling to 8 weeks post-walking onset. As infants continued to progress through their motor milestones, such as transitioning from crawling to walking, there was growth in their receptive vocabulary skills (Walle & Campos, 2014). This effect was replicated in a cross-sectional sample of walkers and non-walkers, such that 12-month-old walking infants had greater vocabulary sizes than the 12-month-old crawling infants (Walle & Campos, 2014). Additionally, the positive, correlational relationship between motor and language skills was also observed cross-culturally in a sample of Chinese infants (He, Walle, & Campos, 2015). The presence of this relationship within two cultures that have inherently different parenting practices demonstrates the strength of the correlational relationship between locomotor and language development.

Research examining the onset of sitting, again, found a positive relationship between motor skills and receptive vocabulary size (Libertus & Violi, 2016). The onset of sitting is crucial because it allows for free use of the infant's hands, as the hands are no longer needed to support the infants' posture. Independent posture allows for greater object exploration, therefore

greater learning about their environment through the exploration of object properties (Rochat & Goubet, 1995). A shift in locomotor skills between crawling and walking also relates to a shift in environmental exploration—walking infants cover more ground in their environment (Thurman & Corbetta, 2017). Therefore, with this gain in locomotor skills, there are subsequent gains in access to objects and components of their surroundings to explore.

The previously described research demonstrates a robust, correlational relationship between the two areas of motor and language development. However, the factors influencing this relationship are not yet fully understood. A similar longitudinal study followed infants from pre-crawling (but sitting independently) into the second year of life; these results showed a similar trend of relationship between motor and language development (Oudgenoeg-Paz, Leseman, & Volman, 2015). Importantly, environmental exploration was found to be a mediator of the relationship between early motor skill onset and greater spatial language (words such as push or climb). The findings from the studies described thus far in this section highlight the importance of exploration in promoting language development. Furthermore, they emphasize the role of the environment in the early stages of language acquisition. These findings reveal interesting relationships between motor skills and other domains of development, yet they only look at the outcome of the relationship. What remains unanswered in this subfield is how the relationship between these two areas jointly impacts the other. Therefore, we need to understand the processes of development.

A classical view of the relationship between motor and language development is that the changes seen are purely a result of general maturation; those that mature earlier in motor skills will see early maturation in language skills as well (Gesell, Thompson, & Amatruda, 1934). However, for this claim to be valid, then the relationship between motor and language

development would have to be bi-directional, such that early language development would predict more mature motor skills later in development (Libertus & Violi, 2016). In a cohort of older infants (ages 18 to 36 months), greater motor skills at 18 months predicted language at 36 months; however, the bi-directional relationship was not present (Wang, Lekhal, Aarø, & Schjølberg, 2014). Interestingly, language skills at 18 months did not predict motor skills at 36 months.

The lack of bi-directionality between the age of onset of motor skills and following language skills indicates the potential for mitigating factors beyond upwards biological maturation. A couple of factors could moderate the relationship. A proposed explanation comes from the social perspective; a shifting language environment in which caregivers interact with their infants differently during different motor milestones. The traditional assumption is that walking infants experience greater language input due to their greater environmental exploration. In one case, however, crawling and walking infants experienced similar language environments (Walle & Campos, 2014). In another case, crawling infants received more adult words than walking infants (Oudgenoeg-Paz, 2014). These trends show that there is something more to the relationship between motor and language development besides shifting language environments.

An alternative hypothesis to the interpretation of general maturation and shifting language environments is that the onset of motor skills begins an unfolding developmental cascade (Bertenthal, Campos, & Barrett, 1984; Libertus & Violi, 2016). As infants see gains in their motor skills, there is an uptick in interactive bouts in their environment (Thurman & Corbetta, 2017). Increases in interactive bouts relate to increased perceptual associations within the environment due to looping perception-action cycles (Corbetta, DiMercurio, Wiener,

Connell, & Clark, 2018). The current proposal of an integrative approach to language development centers around the relationship between increased perceptual associations gained through manual exploration and subsequent object-label mapping. Therefore the developmental cascade that begins with the onset of motor skills could be of importance to language acquisition.

Object Exploration and Object-Label Mapping

The previously described work has shown that object and environmental exploration is important in the development of later language skills. However, these tasks have centered around measuring receptive and productive language skills, not the task demands behind learning novel object-label maps. Research examining the role of object interaction and novel object-label mapping in a semi-naturalistic free-play task had found that an increase in the amount of object manipulation facilitated increased learning, but only when the label occurred during manipulation (Yu & Smith, 2012). Infants in this task had the opportunity to manipulate and explore three different objects on a table, while their caregiver freely labeled the objects using assigned novel words. Caregivers were not given directions regarding how often to produce the novel labels, nor when to produce the labels, they were only given instructions as to which object belonged to which word. These findings were later replicated using a similar paradigm (Pereira, Smith, & Yu, 2014).

The key tenant to these studies is how object exploration facilitates changes in visual attention towards the object, therefore changes in cognition. Each member of the caregiver-infant dyad wore a tiny camera on their forehead to record their visual field. It is important to note that this method does not show where the individuals are looking, as individuals can orient their head in one direction, and their eyes in another. When infants were manipulating the

objects, the novel object dominated a vast majority of their visual field, as compared to when adults held the same object (Yu & Smith, 2012). The obscuring effect is due to the constraints of the infants' body; infants' torsos are shorter than adults, therefore closer to the surface of the table. Additionally, infants' arms are shorter, therefore closer to their head and body as well. These bodily constraints create an optimal moment of visual viewing, because the object dominates the majority of their visual field, therefore obscuring potential distractors (Yu & Smith, 2012; Pereira et al., 2014). Previous bodies of literature have claimed that infants experience difficulties learning from their environment due to the clutter and ambiguity of stimuli (Quine, 1964). However, Yu and Smith (2012) demonstrate how infants can overcome this sensory overload to narrow into specific features of their environment and successfully learn a novel word.

While the previous studies provide great insight into novel object-label mapping as it relates to a dyadic free-play session with a familiar caregiver, several factors limit the results. One of the factors that could influence the results is the role of a familiar caregiver in teaching novel labels. An effect previously found involving infants learning to reach successfully in a task was not replicated when the familiar caregiver that greatly supported the infant during the task was replaced with a neutral, non-communicative experimenter (Corbetta, Williams, & Haynes, 2016; Williams, Corbetta, & Guan, 2015). The authors conclude the effect could be seen because of scaffolding a caregiver can provide, therefore obscuring the effects of the desired manipulation. While an approach that uses a caregiver provides high naturalistic validity, it could misrepresent the finer interactions impacting the task. Another factor, which Yu and Smith (2012) and Pereira and colleagues (2014) discuss, is that the results are correlational as there was no direct manipulation of conditions. The free-play nature of this task

did not allow for proper control of the timing of the novel label production, nor did it for when or if the infant could manipulate the objects. The studies do show the positive relationship between manipulation and object-label mapping, therefore this is a fruitful area of the literature that can be better explained with direct manipulations of the effects of timing and object manipulation, which the present study seeks to do. Understanding the relationship between object interaction in a task, and overall environmental exploration is important because a standard paradigm in language research involves 2-D screens. If object interaction facilitates object-label mapping, which is typical in infants' environments, then studies involving only visual moments of learning on a screen could potentially be missing potential boosts to mapping from manual interaction.

Outside of the literature on cognition, there is also a growing body of evidence for the relationship between environmental exploration and language development in the play literature. This subset of literature focuses mostly on the social aspect of play and learning between the learner and an older caregiver or partner in the interaction. When looking at the dyadic nature of language development, infants are more likely to learn novel words when they are engaged in a one-on-one play session with a caregiver, as opposed to learning the word outside of a play session (Baldwin et al., 1996). This perspective goes beyond the social-pragmatic view because it integrates the factor of object interaction and environmental exploration. Consistent with Yu and Smith's (2012) findings, infant guided interest plays a significant role in object-label mapping during free-play sessions (Dunham, Dunham, & Curwin, 1993). Toddlers were more likely to learn a word if their play partner followed the toddlers' interest rather than trying to pull the toddlers' interest towards another labeled object (Dunham, Dunham, & Curwin, 1993).

Components of Task Performance Success

Several task dynamics could be interacting to potentially impact novel object-label mapping including; attention, synchrony, and dynamic visual experiences. Each of these is inherent in an object-label mapping task that involves manipulation. However, components are also seen in traditional tasks that do not involve object manipulation. This section breaks down the relationship between the different components of the study from Yu and Smith (2012) and relates them to how infants may be learning labels in the real world.

Attention. One of the major explanations of the trends seen in Yu and Smith (2012) is a term they coined embodied attention. Embodiment, in general, is physical interactions with the environment (Thelen et al., 2001), and embodied attention arises explicitly due to an individual's physical interactions with the environment (Yu & Smith, 2012). In this case, the infant's actions on the object lead to an optimal visual moment or increased attention towards the object. Yu and Smith proposed it was the embodied attention that was related to successful object-label mapping. What is not clear, however, is if it is the action itself (the exploration) that led to increased attention, or if it was the single view moment of the object, which facilitated attention. Embodied attention highlights that these two interpretations are dynamically related. Embodied attention alone, however, cannot separate the details of manipulation, such as inferring if it is motor interaction significantly impacting object-label mapping or just attention.

Goal-directed actions in a task can facilitate attention in a task, separate from creating optimal moments of attention. For example, a recent study by Wiener (2016) demonstrated that infants, who disengaged previously with a task, refocused their attention to the task when introduced to the action of reaching to an object within the otherwise same paradigm (Wiener, 2016). The first half of Wiener's (2016) task involved infants only viewing objects, to see their

looking patterns. As the task progressed, infants increasingly looked less towards the object and became more off-task. Halfway through the task, the experimenter provided infants with an opportunity to manipulate the objects by moving the object towards the infant. After the infants had experience manipulating the objects, their looking patterns again shifted to increased looking to the object, as compared to being off-task (Wiener, 2016). These findings demonstrate that introducing goal-directed action in a task, such as reaching for an object, can increase attention to the presented object. Furthermore, when novel objects are presented to infants in periods of high attentiveness, then they are more likely to remember those objects (Frick & Richards, 2001). Infant's action on objects in their environment could relate to increased attention towards those items, which, in turn, could be a mediator of novel object-label mapping within the relationship between motor and language development.

Synchrony. A component of the task from Yu and Smith (2012) that is central to the interpretation is the role of synchrony between hearing and holding and intersensory redundancy in their paradigm. There were several cases of synchrony in this paradigm. When the infant saw the object and manipulated it, but did not hear the novel word label; when the infant saw the object, and heard the novel word label, but was not manipulating it; and finally, when the infant saw the object, heard the novel label, and was manipulating the object all at the same time. The results showed that the production of the novel label, while the infant was holding the object, facilitated better object-label mapping; however, there are multiple reasons as to why this could have occurred. The authors propose that the label was provided during a moment of optimal attention towards the object. However, an equally compelling explanation is that there was increased learning due to the convergence of sensory systems. Redundant information across senses, such as a ball dropping on a hard surface, making a noise when it hits

the table, draws attention towards the perception of the events (Gogate, Bahrick, & Watson, 2000). Infants, in particular, are sensitive to synchronous information, or whether or not object visual properties and sounds are happening in synch (Lewkowicz, 1996).

One of the ways synchrony has been examined with language development is through the social environment of the caregiver-infant dyad. For example, Gogate, Bolzani, & Betancourt (2006) found that when caregivers taught infants novel words mapped onto a novel object, the infants better learned the words if the caregiver physically moved the object in conjunction with the inflection of their voice (Gogate, Bolzani, & Betancourt, 2006). For example, if the word started with a strong inflection followed by a weak inflection, the caregiver would map the height of the object with the inflection of their voice. In this case, we see that one specific manipulation of synchrony can influence object-label mapping. This type of synchrony is difficult to produce and is not too common in natural environments, however. More often, infants will experience synchronous events of labeling while holding, such as the ones seen in Yu and Smith (2012). In a task that involved infants learning to reach, increases in reaching coincided with when the infant experienced a synchronous moment between their haptic, visual, and auditory senses, as compared to just haptic and visual (Williams & Corbetta, 2016). These results demonstrate that infant learning, in this case learning during a reaching task, is better facilitated when haptic, visual, and auditory senses converge. Similarly, I am proposing it was the convergence of those three domains that led to better object-label mapping in the Yu and Smith (2012) paradigm. The present study seeks to expand these findings by taking a finer look at the differing levels of synchrony through a dynamic systems approach.

Dynamic Visual Experience. When infants interact with and manipulate objects, they can move it in a way that creates unique, self-generated object views that they would not

experience if the object was on a screen or held stationary and out of reach. During the period of development when infants experience a vocabulary spurt, there is also a shift in how infants view objects (Pereira, James, Jones, & Smith, 2010). Infants that had the shift towards a preference for planar object views (perpendicular or parallel to the infant's line of sight) also had larger vocabulary sizes at the time of the study (James, Jones, Smith, & Swain, 2014). This creates a link for dynamic object views and language development.

A similar finding has been found longitudinally showing a relationship between dynamic object views and vocabulary size. When infants create more variable object views during manipulation at 15-months-old, they demonstrated larger vocabulary six months later (Slone et al., 2019). Interestingly, caregiver's created object variability did not impact infant vocabulary, indicating that it must be self-generated. Infants' propensity to create dynamic object views could be a contribution to the individual differences seen in language development. Slone and colleagues propose that understanding infants' experiences with objects could begin to explain differences in language growth. Both of these studies link object views and vocabulary size, however, direct associations between dynamic views and learning have not yet been established.

A Dynamic Systems Approach to Language

Within the past decade, there has been a call to put infants back together, termed the "Humpty Dumpty Problem" (Oakes, 2009). This problem refers to the plethora of infant researchers studying specific processes independently, yet we do not have a great understanding of how these processes work together. Often, models of developmental research, such as social, motor, or cognition are not integrative, yet that is how the infant develops—as a whole, cohesive unit. An approach to understanding infant development, and specifically to the co-occurring, interacting development of motor skills and language, is to view development as a dynamic system. Dynamic systems theory offers a different lens on development, such that

facets of development should not be viewed as an independently occurring phenomenon, but rather the product of multiple and continuous interactions within the developing system (Thelen & Smith, 1998).

One of the central tenants of dynamic systems is that development of a system does not necessarily follow a linear trajectory and, importantly, different system levels of the individual are in near constant interaction (Connell, DiMercurio, & Corbetta, 2017). Gershkoff-Stowe (2001) demonstrated an example of the dynamic nature of language with a sudden regression in older infant's ability to map labels onto known words. Infants in this study were presented with familiar, common words such as duck and shoe. When asked to name a familiar object, such as a duck, the infants were successful in naming the stimulus as a duck. However, when another stimulus appeared immediately after, in this case, a shoe, the infants made an error and labeled the shoe as a duck. This error was not always present longitudinally; infants in this study experienced high success in familiar object labeling, followed by the regression of the errors noted previously, and finally returned to a typical success in object-label mapping, showing a "U" shaped developmental trend (Gershkoff-Stowe, 2001). This demonstrates that infants know what the labels are because of their prior success in labeling, yet, the error still occurs. When we consider language to be a dynamic system, there is a clearer explanation of the regression seen in object-label mapping. This example can be related to coupling, a tenant of dynamic systems—at this particular stage of development where the regression occurs, retrieval and prior-activation of a word are tightly coupled, and infants cannot overcome the priming effect of the previous object-label mapping (Gershkoff-Stowe, 2001). As the system continues to develop, retrieval and prior-activation are no longer tightly linked. Therefore the error disappears.

The relationship between motor and language development can also be viewed dynamically. The onset of motor skills allows for increased environmental interaction, which in turn relates to an increase in sensory experiences, therefore learning about the environment via the perception-action loop (Corbetta, DiMercurio, Wiener, Connell, & Clark, 2018; Corbetta & Snapp-Childs, 2009). An increase in environmental interaction is related to increases in attention created by the individual's interaction with the environment, which can facilitate object-label mapping, but only if the timing of the novel stimulus is synchronous with manipulation (Yu & Smith, 2012). The present study aims to test a dynamic systems model of novel object-label mapping by integrating factors of environmental manipulation, attention, and synchrony. The proposed model has several dynamical levels (see Figure 1). Object manipulation could impact novel object-label mapping, by influencing attention to the object during subsequent presentations of the object. Furthermore, the timing of the production of the novel label could impact subsequent object-label mapping in relation to the infant's manipulation of the object, if the label production relates to increased manipulation. Overall, the review of the literature has shown that these factors are all jointly related, yet despite their close connections, attention, environmental exploration, and synchrony have not been systematically assessed in a singular study of object-label mapping. I aim to understand how aspects of motor interaction, through the concurrent and dynamic interaction of seeing while holding and hearing, will relate to gains in conceptual knowledge, through the specific example of novel object-label mapping.

Purpose and Hypotheses

The present study seeks to understand the degree to which the effects of manipulation of objects in a novel object-label mapping task, as well as the timing of the production of the novel

word label, influences subsequent learning. Additionally, the present study aims to assess if there are individual differences in how infants learn novel object-label mappings. To test the above questions, the present study had three conditions all within subjects (Look-Only, In-Sync, and Out-of-Sync). These conditions will be described in greater detail during the methods sections, but it is important to note that they differ in whether the infants can manipulate the target object and in when the novel label is provided. The Look-Only condition involves no manipulation. The In-Sync condition involves label production synchronous to the infant's manipulation, while the Out-of-Sync condition has the label production asynchronous to the infant's manipulation. The Look-Only condition will serve as our control condition as it is similar to how many typical object-label mapping studies occur, albeit with a 3-D object, and not on a screen. Infants also cannot interact with every object they encounter in their environment, so this condition provides a similar experience to how many infants receive novel labels.

If infants are sensitive to optimal moments of attention through hearing the label while holding and seeing the object, then the In-Sync condition, where these three modalities converge, should produce the highest rates of object-label mapping. Furthermore, because there are individual differences and variability seen in language outcomes, such as vocabulary scores, then there should also be variability in how infants are learning how to map labels onto novel objects. Therefore, if synchrony between seeing the object while holding it and hearing the novel label is supposed to be the best condition for learning, then infants who initially learn in the more theoretically difficult, no manipulation condition, should learn across all three conditions. However, if infants struggle with the Look-Only condition, they should see a substantial boost in learning when the label is synchronous to manipulation.

An alternative hypothesis to the prior predictions is that manipulation will not predict learning; rather learning will be reflective of attention towards the target object during the presentation trials. Embodied attention posits that manipulation is essential because of the singular object views that the infant creates, therefore an optimal moment of attention (Yu and Smith, 2012). Therefore, if attention alone impacts novel object-label mapping, then there should not be a difference in condition performance as long as the infant attended to each target object, which is presented in an uncluttered scene, equally.

CHAPTER THREE: METHOD

Participants

Thirty infants, aged 18- to 21-months old ($M = 19.4\text{mo}$, 17 females, 13 males), participated in this study.

Stimuli and Apparatus

The infants were seated in a standard highchair using a 7-point strap that secured the shoulders, chest, and torso of the infant. This assured that the infants would safely remain in the seat and could not lean too far forward. The highchair was set in front of a tri-fold, solid black theater that had an open window with a curtain in the middle. The experimenter was seated behind the theater to facilitate the object presentation but was obscured by a curtain. A computer monitor was fitted into the open window of the theater initially for calibration of the eye-tracker but was removed for the object presentations. A Tobii X2-60 eye-tracker (collecting at 60Hz) rested on the theater, 60 cm across from the infant, with the angle positioned up towards the infant's eyes. Tobii Pro Studio was used to collect the data during the sessions. Figure 2 shows the experimental setup.

A webcam was affixed to the front of the theater and adjusted to capture the infants' face and upper torso so that the object manipulation could be later coded. There were also two video cameras (Panasonic PV-GS39) on tripods on either side of the infant. The video fed into a Digital Video Switcher SE-500 (Datavideo Corp., Whittier, CA, United States) where the image was recorded and split for a side-by-side view of the infant. This allowed for each side of the infant to be recorded so that behavior and object manipulation could later be analyzed. An additional camera was placed to the left side of the infant to show the infant's behavior to an obscured person running the eye-tracker; this video was not recorded.

There were six objects used in this study, as seen in Figure 3. All objects had novel shapes and were made of wood. The objects were sanded, painted with non-toxic acrylic paint, and then sealed with a water-proof, non-toxic sealant layer. Each object was painted a different bright and stimulating color. Three of the objects were paired with a novel word (*dobu*, *nilla*, *modi*), which all have a similar phonotactic probability ($M = 1.18$). Therefore they share similar chances of occurring in the English language. Each participant was presented with three objects paired with a novel label and three non-labeled objects, and all conditions were within subjects.

Auditory stimuli were prerecorded and played during the session during the presentation phase and the test trial phase. Each object had an auditory stimulus that was played over speakers oriented towards the infant during each presentation trial. The recordings were the same for each object, with the only variation being the novel label used, which repeated three times with a carrier phrase and once in isolation: “Look at the *modi*! There’s the *modi*! It’s a *modi*! *Modi*!”. The non-labeled objects were paired with a neutral auditory stimulus: “Look at this one! Here it is! Do you see it?”. The test trial stimuli differed in the production of the target label, which was produced three times with a carrier phrase and once in isolation for each test trial: e.g., “Where’s the *modi*? *Modi*! Can you find it?” followed by “Find the *modi*! Get the *modi*!”. There were approximately 1500ms of silence while the objects were in view, before the start of the auditory test stimulus, to gather baseline looking data. The onset of the test label occurred at 2000ms. The second label began at 4500ms. The whole string lasted approximately 6800ms, see Figure 4 for an example of the timeline.

Questionnaire and Vocabulary Assessment

During the session, questionnaires were given to the caregiver that included demographic information and self-report measures of the infant’s onset of walking and talking. Due to the

nature of the self-report, we expect these onset estimates to be weak. Additionally, caregivers were given a short-form, Level 2 MacArthur-Bates Communicative Development Inventory (MCDI) (Fenson, Marchman, Thal, Dale, & Reznick, 2014). The MCDI had 100 words and parents were instructed to indicate the words that their infant could produce.

Procedure

Each participant came into the lab for a one time visit lasting 30 to 45 minutes. The infants proceeded through a presentation phase and a testing phase. Each of the three labeled objects was randomly assigned to one of three presentation conditions, designed to test the effects of the object manipulation and label synchrony on label learning (Look-Only condition, Out-of-Synch condition, and In-Synch condition). Since the hypothesis is that type of presentation condition would result in differing levels of learning (i.e., a condition with manipulation should be better learned than a condition with no manipulation), the labeled objects were never tested against each other. The infant would probably default to the object most likely to be learned, rather than selecting based on what they know. For this reason, there were unlabeled objects matched to each labeled object and they received the same condition exposure.

After obtaining consent from the caregiver, infants were securely fastened into the highchair. Initially, the computer monitor was in place for the calibration process. A friendly children cartoon played on the screen to capture the infants' attention and orient them to the center of the theater. This orienting allowed the experimenter to view the infants' gaze on the eye-tracker and to see if any physical angle adjustments needed to be done to capture the infants' gaze better. Once the experimenter was satisfied with the positioning of the infants' eyes, calibration began. Calibration took place using the computer screen that is in the open

window of the theater. A cartoon animal paired with a noise flashed on the screen in 5 different points (four corners and the middle), to orient the infants' gaze towards the point. When the infants looked at the point, the experimenter proceeded to the next point. Once all 5 points were obtained with high accuracy and little noise, the screen was removed, and the presentation trials began.

Each of the infants proceeded through 15 presentation trials and 12 test trials. The presentation trials were pseudo-randomized into four blocks. Each block consisted of one presentation trial of each labeled object that was randomly ordered within the block. This was done to ensure that each label was potentially learned at a distributed pace, and the infants did not receive the same label twice in a row. The three non-labeled objects were presented one time each throughout the presentation trials. The non-labeled objects were presented only one time each, using their matched condition procedure. This was done to reduce the number of trials the infants would need to endure before moving into the test trial phase of the study. All objects were presented in the open window of the theater. The eye-tracker and theater for object presentation were 60 cm away from the infants. Therefore, these objects were out of reach to allow for initial visual scanning without manual manipulation.

This task utilized three, within subjects, conditions. The differences between the conditions are described below (see Figure 5 for a summary). However, each condition followed a basic structure containing two phases. First, the object is presented out of reach in a singular view. The second phase involves either infant manipulation or experimenter manipulation of the presented object. For two conditions, the Look-Only and Out-of-Sync, the presentation auditory stimulus was played while the object was in that first, out of reach phase. The In-Sync condition, however, had the presentation auditory stimulus during the second, manipulation

phase. The Look-Only condition used experimenter manipulation, while the Out-of-Sync and In-Sync conditions contained infant manipulation.

Look-Only Condition

Each condition began with an object obscured by a curtain and resting on a tray. At the start of the trial, the curtain was lifted to reveal an object sitting on the surface of the theater. After the infant's first fixation onto the object, the auditory stimuli played, lasting approximately seven seconds. A second experimenter ran the eye tracker and had a live view of the infant, as well as live tracking. Therefore, they were accurately able to begin the auditory string with the first fixation. After the auditory stimulus concluded, the experimenter picked up the object and slowly rotated the object for approximately seven seconds. After seven seconds passed, the trial concluded, and the curtain was lowered.

Out-of-Sync Condition

The timing of the production of the auditory label was similar to the Look-Only condition; infants were presented with the object and labeling phrase in the same manner and with the same timing as in the look-only condition. However, after the auditory stimulus concluded, the experimenter pushed a drawer-like mechanism, on which the object was resting, forward into the infant's reaching space to allow for direct object manipulation. After about seven seconds of object manipulation, the experimenter cued the caregiver to remove the object from the infant which concluded the trial.

In-Sync Condition

The timing of the auditory label differed in this condition as compared to the previously described two conditions. However, the degree of object manipulation is the same as the Out-of-Synch condition. The curtain lifted to indicate the start of the trial. After approximately seven

seconds of accumulated looking time onto the scene, the experimenter pushed the drawer mechanism into the infant's reaching space. The second experimenter, using the live view of the infant on a monitor, was able to determine when the infant first contacted the toy and triggered the start of the auditory stimulus, lasting approximately 7 seconds. After the end of the auditory stimulus, the experimenter cued the caregiver to remove the object from the infant, concluding the trial.

Test Trials

Each labeled object was the target of a test trial four times. The test trials consisted of the target object resting on one side of the tray and the condition-matched but non-labeled object resting on the other side of the tray. The side of the tray for the target object was counterbalanced and pseudo-randomized (repetitions more than two sides in a row were adjusted) for each participant and across all conditions. For the first part of the test trial, the auditory stimulus played while the objects were presented in the open window of the theater, out of reach. After the infants had a chance to search for the target object visually, the rest of the auditory string played, prompting a physical selection of the target object while the two objects were moved forward into the infants' reaching space. The test trial concluded after the infant made a selection from the two objects, or grabbed both, and the caregiver was instructed to remove the object(s) from the infant.

Coding and Analyses

All data were coded using data video coding software Datavyu v1.2 (Datavyu Team, Databrary Project, New York University). All trials were hand-coded. For the presentation trials, an area of interest for the eye-tracking was established allowing an approximate inch around the object. The primary coder coded all of the presentation sessions. In order to establish

reliability, 20% of the trials were randomly selected to be recoded by a secondary coder. Our presentation coding reached 97% reliability for looking on the target object (Cohen's $\kappa = .892$, $p < .0001$). If the fixation fell within the window, then it was considered on the stimulus. The manipulation during presentation trials was also coded. For the In-sync and Out-of-Sync conditions, the primary coder determined when the infant held the object and if the infant was looking at the object. This looking data was coded by looking at the front camera view of the infants' face and their gaze direction. The holding data reached 95.7% reliability (Cohen's $\kappa = .846$, $p < .0001$) and the looking reached 97.4% reliable (Cohen's $\kappa = .663$, $p < .0001$). The Look Only condition involved the experimenter manipulating the object after the initial presentation. The primary coder assessed if the infant was looking at the target object during the experimenter manipulation. The reliability for looking at the target was 88% (Cohen's $\kappa = .648$, $p < .0001$) and 89% (Cohen's $\kappa = .662$, $p < .0001$) for Elsewhere for the Look-Only experimenter manipulation.

For coding the test trials, the scene was split using a midline between the target and distractor objects, as well as a standardized AOI on the upper, bottom, and outer sides of the objects. The location of the fixation was used to determine where the infant was looking. The reliabilities for the test trials for 20% of the data was 95% for the distractor objects and 95% for the target objects (distractor Cohen's $\kappa = .9$, $p < .0001$; target Cohen's $\kappa = .707$, $p < .0001$). Additionally, all fixations that did not fall within the AOIs were coded as "elsewhere". The elsewhere codes were combined in the final analyses with all off-task durations.

Due to the nature of the methodology, where the unlabeled objects were not presented as often as the labeled objects, I expected a slight preference for the unlabeled objects. Figure 4 shows a breakdown of the test string, with a 2000 ms baseline period to gather looking data between the two objects, followed by a 2000 ms window of analysis following the onset of the

first label. Learning was determined as a proportion change in looking to the target object from the baseline window to the window of analysis. A positive value, demonstrating an increase in looking to the target object from the baseline window to the analysis window, indicated successful object-label mapping. A negative value, therefore a decrease in target looking from the baseline window to the analysis window, indicated unsuccessful object-label mapping. Physical selection during the test trials was also coded offline in Microsoft Excel. The selection was considered as the final hold of either the target or the distractor object. The target action of the final hold was coded because initially infants might have touched one object but selected another to pick up. This method of coding the test trial performance was not used in the final analyses as the infants selected both objects on 75% of the test trials. This does not indicate a lack of learning, but likely a lack of inhibition to not grab both objects.

To assess the variability in the learning process, two groups within the sample were created post-hoc. The division of the groups was based on their performance during the Look-Only condition. If the infants did not demonstrate successful object-label mapping, they were put into a Low Look-Only Performance group. If the infants did demonstrate learning, they were put into a High Look-Only Performance group. This split was used to see whether infants who are struggling with conventional learning measures (visual and auditory only) would perform better when their learning environment is shifted to a more dynamic, redundant, and synchronous event (visual, auditory, and haptic). This allowed for a more in-depth look at possible learning processes impacting a foundational skill of language development.

Data for presentation dynamics and test trial performance were analyzed using a Generalized Estimating Equation (GEE). This type of analysis is ideal for clustered and correlated data. However, it is specifically beneficial for analyses on non-normalized data sets (Ghisletta & Spini, 2004). I used the model-based estimator, with an unstructured correlation

matrix on a linear model. The predictor variables were the learning group and condition. The dependent measure for the presentation trials included the proportion of looking time towards the object, task synchrony, manipulation, and looking at the object during manipulation. For the test trials, the dependent measure was the proportion increase of target looking from the baseline to the window of analysis. Lastly, I compared task dynamics from the presentation phase of the study to infants' test trial performance.

CHAPTER FOUR: RESULTS

To assess variability and individual differences, the sample was split based on their Look-Only performance. Learning was assessed as a proportion increase or decrease of looking to the target object from the baseline to the analysis window of the test string. If the proportion of looking towards the target object in the Look-Only condition was negative, indicating no learning, the infants fell into the Low LO (Look-Only) Performer group. If the proportion was positive, indicating learning, then the infants fell into the High LO (Look-Only) Performer group. The membership for the groups was 14 infants in the Low LO Performer group and 16 infants in the High LO Performer group. All presentation task data were ran using a GEE, with the LO performance group and condition as the predictor variables.

This analysis was used in an attempt to control for either novelty or familiarity preferences. There was not an overall preference, but there were many individual differences between having a novelty or familiarity preference. Overall, very few infants demonstrated even looking between target and distractor objects during the baseline window, see Figure 6. A Wilcoxon Signed-Ranks test revealed no difference between the performance group for their baseline looking preference ($Z = -1.036, p = .300$).

Presentation Trial Performance

All infants in the data set successfully processed through the presentation phase. This included four trials for each labeled object (12 trials) and one trial for each unlabeled object (three trials) for every session. The analyses below are on the labeled object trials only. There were 360 trials for the analyses; however, 14 trials were excluded due to gaze loss from the eye-tracker. Therefore, 346 trials were analyzed.

Proportion of Looking During Presentation Trials

The GEE revealed no main effect for the LO Performance groups (Wald $\chi^2(1) = 1.483$, $p = .223$; High LO Performance $M = .507$, $SE = .034$; Low LO Performance $M = .446$, $SE = .037$) nor for condition (Wald $\chi^2(2) = .952$, $p = .621$; Look-Only $M = .461$, $SE = .031$; In-Sync $M = .477$, $SE = .031$; Out-of-Sync $M = .49$, $SE = .031$). There was a marginal effect for LO Performance group by Condition interaction (Wald $\chi^2(2) = 5.777$, $p = .059$), however, none of the pairwise comparisons were significant (see Table 1 for means). Thus, looking towards the stimulus did not differ between LO Performance groups nor condition.

Looking While Hearing During Presentation Trials

Each participant received four repetitions of each label during each trial, resulting in exposure to the label 16 times per condition. This analysis examines beyond looking to the stimulus during the presentation phase, but assesses if the infant was attending to the object while hearing the label.

For the Look-Only and Out-of-Sync conditions, this entailed fixating onto the object while hearing the label. For the In-Sync condition, looking at the object while hearing the label also entailed holding the object at the same time. The GEE showed no main effect for group membership (Wald $\chi^2(1) = .684$, $p = .408$). However, there was a main effect of looking while hearing between the conditions (Wald $\chi^2(2) = 23.934$, $p < .0001$, see Figure 7). There was more looking while hearing in the Look-Only condition ($M = .521$, $SE = .035$) than the In-Sync ($M = .36$, $SE = .035$) and Out-of-Sync ($M = .55$, $SE = .035$, see Table 2) conditions ($ps < .0001$). The In-Sync and Out-of-Sync conditions did not differ from each other ($p = 1.00$).

The GEE also revealed a LO Performance Group by Condition interaction (Wald $\chi^2(2) = .6122$, $p = .045$, see Figure 8). The Low LO Performers showed a difference looking while

hearing between In-Sync ($M = .36$, $SE = .051$, see Table 3) and Out-of-Sync ($M = .567$, $SE = .051$, $p = .039$). The High LO Performers differed in their looking while hearing between the Look-Only ($M = .598$, $SE = .048$) and In-Sync ($M = .36$, $SE = .048$) conditions ($p < .0001$). Between LO Performance groups, the Low LO Performers looked to the object while hearing the label less frequently in the In-Sync condition than the High LO Performers were in the Look-Only condition ($p = .01$). There were no differences related to the High LO Performer's Out-of-Sync condition ($M = .534$, $SE = .048$). Overall, there were no differences between LO Performance groups within each condition, showing that they looked to the object while hearing the label the same amount for each condition. However, the LO performance groups revealed differences between conditions.

Manipulation During Presentation Trials

These analyses are on the proportion of time the infant were holding the object during the manipulation phase of presentation trials. Therefore, these analyses were only run on the In-Sync and Out-of-Sync conditions. A GEE showed no main effect for the LO Performance groups (Wald $\chi^2(1) = 1.725$, $p = .189$. There was a main effect for condition (Wald $\chi^2(2) = 5.53$, $p = .019$, see Figure 9), however, there was not a LO Performance group by condition interaction (Wald $\chi^2(2) = 2.968$, $p = .085$). The amount of time infants could hold the object was equal across the two conditions, yet infants held the toy significantly more in the In-Sync condition than the Out-of-Sync condition (In-Sync $M = .941$, $SE = .027$; Out-of-Sync $M = .885$, $SE = .027$, see Table 4).

Manipulation While Looking During Presentation Trials

Infants either experienced observed manipulation (Look-Only) or manual manipulation (In-Sync and Out-of-Sync). These analyses compare the proportion of time the infants were

looking at the object either while observing the experimenter manipulate the object, or while the infants were manipulating the objects. A GEE showed no difference between LO Performance groups (Wald $\chi^2(1) = 2.048, p = .152$). Therefore they looked at the object during manipulation equally. There was a main effect of condition (Wald $\chi^2(2) = 70.207, p < .0001$, see Figure 10), however, there was not a LO Performance group by condition interaction (Wald $\chi^2(2) = 3.98, p = .137$). Infants looked at the object during manipulation more when it was the experimenter moving the object, more than both conditions involving manual manipulation ($ps < .0001$; Look-Only $M = .615, SE = .031$; In-Sync $M = .338, SE = .031$; Out-of-Sync $M = .342, SE = .031$, see Table 5). There was no difference between the In-Sync and Out-of-Sync conditions ($ps = 1.00$).

Performance During Test Trials

All infants but two completed all 12 test trials. The two infants that did not complete all 12 test trials missed up to two test trials. This left 357 trials for analyses. Of these 357, 20 trials did not have sufficient eye tracking, indicating gaze loss. Additionally, there were exclusion criteria for the test trials such that if the infants did not attend to either object during the windows of analysis, then they were excluded from the analyses. Sixty-four trials were excluded. With both the insufficient eye tracking and no object looking exclusions combined, 273 test trials remained for analyses. For all analyses, the data were not normalized; thus, nonparametric tests were used. The test trial analyses were run using a GEE. Therefore, all reported means are estimated marginal means with their standard errors.

Performance Between the Learning Groups

The predictor variables for the GEE were LO Performance Group and condition, with the dependent measure of proportion change in looking from a 2000ms baseline window to a 2000 ms window of analysis that occurred following the onset of the first label. The main

effects are LO performance group and condition. Between the two LO Performance groups the GEE showed no difference in learning across the three conditions (Wald $\chi^2(1) = 3.125, p = .077$; High LO Performer $M = .038, SE = .03$ Low LO Performer $M = -.046, SE = .033$). Overall, there was not a difference in learning between the conditions (Wald $\chi^2(2) = .633, p = .729$; Look-Only $M = -.01, SE = .051$; In-Sync $M = .03, SE = .0521$; Out-of-Sync $M = -.039, SE = .053$). This shows that test trial performance was not based on either the LO Performance group or condition alone.

The GEE revealed an interaction between the High LO Performer and Low LO Performer infants and their condition performance (Wald $\chi^2(2) = 15.375, p < .0001$; see Figure 11). The Look-Only condition between the High LO Performer and the Low LO Performer infants during the test differed significantly ($p = .002$; High LO Performer $M = .183, SE = .069$; Low LO Performer $M = -.203, SE = .074$, see Table 6). Within the Low LO Performer infants, the Look-Only condition differed from the In-Sync condition, where the infants performed significantly better in the In-Sync condition ($p = .044$; Look-Only $M = -.203, SE = .074$; In-Sync $M = .120, SE = .076$, see Table 2). For a full breakdown of all the insignificant pairwise comparisons see Table 7. These findings demonstrate that different conditions benefited differently performing infants (see Figure 12 for a breakdown of the individual differences for test trial performance).

Study Dynamics and Performance

Age and Performance

There was not a significant difference in ages between the two learning groups as shown by a Wilcoxon Signed-Ranks test ($Z = -5.25, p = .60$; High LO Performer $M = 19.506, SD = .652$; Low LO Performer $M = 19.429, SD = .565$). A Pearson correlation showed no impact of age on test performance in the Low LO Performer group (Look-Only $r(14) = -.217, p = .456$; In-

Sync $r(13) = .142, p = .643$ Out-of-Sync $r(12) = .122, p = .707$). Similarly for the High LO Performer group, age did not impact learning (Look-Only $r(16) = .434, p = .093$; In-Sync $r(15) = -.076, p = .786$ Out-of-Sync $r(15) = -.460, p = .085$). Thus, age was not a factor in either group.

MCDI and Performance

A Wilcoxon Signed-Ranks test showed that the infants' MCDI scores did not significantly differ between the two learning groups ($Z = -.280, p = .780$; High LO Performer $M = 23.938, SD = 19.519$; Low LO Performer $M = 21.214, SD = 23.538$). A Pearson correlation revealed there was no evidence for a relationship between MCDI and test performance for Low LO Performers in any of the three conditions (Look-Only $r(14) = .013, p = .965$; In-Sync $r(13) = .357, p = .231$ Out-of-Sync $r(12) = .134, p = .679$). There was no impact of MCDI on performance in the High LO Performer group either (Look-Only $r(16) = .232, p = .386$; In-Sync $r(15) = -.320, p = .244$ Out-of-Sync $r(15) = -.284, p = .306$). In sum, there was no evidence for MCDI influencing the infants' task performance.

Walking Onset and Performance

Walking onset was self-reported by the caregiver at the session. A Wilcoxon Signed-Ranks test showed no differences between the two learning groups ($Z = -1.741, p = .082$; High LO Performer $M = 13.531, SD = 1.477$; Low LO Performer $M = 12.167, SD = 2.082$). There was no significant impact of walking onset between condition performance between the Low LO Performer group (Look-Only $r(12) = -.056, p = .862$; In-Sync $r(11) = .077, p = .822$ Out-of-Sync $r(11) = .317, p = .342$) and the High LO Performer Group (Look-Only $r(16) = .188, p = .485$; In-Sync $r(15) = -.173, p = .536$ Out-of-Sync $r(15) = .303, p = .272$). Hence, there is no evidence for an effect of onset motor milestone for these infants.

Talking Onset and Performance

Talking onset was self-reported by the caregiver at the session. A Wilcoxon Signed-Ranks test showed no difference between the two learning groups ($Z = -.510, p = .610$; High LO Performer $M = 8.967, SD = 2.635$; Low LO Performer $M = 10.273, SD = 2.524$). A Pearson correlation showed no relationship between onset of talking and test performance in either of the learning groups (Low LO Performer: Look-Only $r(11) = -.206, p = .543$; In-Sync $r(10) = .433, p = .212$ Out-of-Sync $r(10) = -.178, p = .622$; High LO Performer: Look-Only $r(15) = .187, p = .506$; In-Sync $r(15) = .015, p = .958$ Out-of-Sync $r(14) = -.101, p = .732$). Therefore, infants' talking onset did not impact test trial performance.

Looking During Presentation and Performance

A Pearson correlation revealed no significant relationship between time spent looking towards the object during presentation and test trial performance for the Low LO Performer group (Look-Only $r(14) = -.121, p = .681$; In-Sync $r(13) = -.513, p = .073$; Out-of-Sync $r(12) = -.322, p = .307$). Similarly, there was no correlation within the High LO Performer group (Look-Only $r(16) = .097, p = .722$; In-Sync $r(15) = .137, p = .627$; Out-of-Sync $r(15) = -.196, p = .485$). This means that the amount of looking to the object did not relate to test trial performance.

Looking and Hearing During Presentation And Performance

The frequency of looking to the object while hearing the label during the presentation trials did not have a significant relationship with test trial performance, as indicated by a Pearson correlation, in the Low LO Performer group (Look-Only $r(14) = -.151, p = .529$; In-Sync $r(13) = -.096, p = .756$; Out-of-Sync $r(12) = .092, p = .776$) and the High LO Performer group (Look-Only $r(16) = .106, p = .697$; In-Sync $r(15) = .240, p = .390$; Out-of-Sync $r(15) = -.037, p = .896$). Therefore, the amount of synchrony did not influence object-label mapping.

Manipulation During Presentation and Performance

A Pearson correlation revealed that the proportion of time infants held the object during manipulation did not significantly impact performance in either the Low LO Performer group (In-Sync $r(13) = .108$, $p = .726$; Out-of-Sync $r(12) = -.145$, $p = .652$) and the High LO Performer group (In-Sync $r(15) = .085$, $p = .763$; Out-of-Sync $r(15) = -.183$, $p = .514$). Thus, how much the infants manipulated the object did not impact test trial performance.

CHAPTER FIVE: DISCUSSION

The present study provides evidence that dynamic synchrony between looking and holding an object, while simultaneously hearing a label can provide a substantial boost in novel object-label mapping for infants that struggle with typical visual only learning experiences. These findings reveal the prevalence of individual differences in early language learning that is not predicted by common factors, such as MCDI scores, age, and locomotion. The lack of significant differences between common factors indicates that different learning experiences may benefit infants depending on their individual differences. These results call for a more in-depth exploration into the different task dynamics that may be at play while learning a word.

A Dynamic View of Learning

Interestingly, for the LO Performance groups, there was not an effect of common factors that typically account for individual differences such as motor milestones or MCDI. These factors did not impact test trial performance, and there was no difference between the two LO Performance groups based on individual learning variations. Importantly, there was not a factor of age and performance, therefore the present study does not support a maturationist view positioning growth with better language skills. The differences seen in this study stems from the impact of task dynamics of the different conditions.

The performance trial results provides an interesting addition to the data. Throughout all analyses, there was not a main-effect of LO Performance group, showing that the infants, regardless of group, attended to the task the same way. There is not a difference in the amount of looking, looking while hearing, manipulating, or looking during manipulation. Therefore, these data support that different task dynamics influence learning differently across infants. There were differences in how infants, regardless of LO Performance group, approached the task between conditions. Regardless of LO performance group, infants held the object for a

significantly greater amount of time in the In-Sync condition than in the Out-of-Sync condition. Although the holding time was high for both conditions, the difference demonstrates that label timing can feed into and influence manipulation. The amount of holding could have highly benefited the low LO Performing infants. Additionally, the looking time data for the presentation trial justifies that attention to a stimulus is not enough to drive object-label mapping for some infants. Infants attended to the objects equally across conditions and between LO Performance groups. If attention to a stimulus alone were enough to drive learning, then there would have been learning across all conditions due to the equality of looking during presentation trials.

The Look-Only condition is representative of traditional tasks that measure language learning, and is one of the ways objects are commonly labeled in an infant's environment. Yet, infants performed fundamentally differently in the Look Only condition as compared to the other two conditions. Nearly 50% of the infants demonstrated no learning in the Look Only condition, which is substantially higher than in the two other conditions. This study can begin to address why there is such variability in a typical mapping task. Some infants need to have intermodal synchrony between hearing, holding, and seeing to map a novel label onto an object successfully. Furthermore, these results provide support for the proposed dynamic model of novel-object label mapping (see Figure 1), specifically for the infants that struggle with visual only experiences. The convergence of these attention, manipulation, and label timing from the dynamic model greatly benefited the Low LO Performing infants who did not learn in a visual-only condition.

The Role of Holding while Hearing

The lack of significance in the amount of looking while hearing between the three conditions as it relates to subsequent object-label mapping indicates that the amount of

synchrony of looking while hearing does not impact learning, but that it is rather the presence of the looking while hearing and holding for the Low LO Performer group. Therefore, the quality of the synchronous events is more important than the quantity for the Low LO Performing infants. Interestingly, rates of looking while hearing were relatively low for the In-Sync condition: out of the twelve repetitions of the word across the four trials, infants were only looking and hearing while holding on a third of the label productions. Additionally, infants from the Low LO Performer group looked to the object while hearing the label far more during the Look-Only condition than the In-Sync condition. Despite that low amount, there was still a significant benefit for the Low LO Performer group. Although there was more looking while hearing, it did not provide the same benefit of looking and holding while hearing.

To understand the interaction between the High LO Performer and Low LO Performer groups, we have to consider the dynamics in each condition. Throughout all conditions, infants experienced object view variability: It was either self-generated (In-Sync and Out-of-Sync) or created by the experimenter (Look-Only) through manipulation. Although Low LO infants did not differ in the amount of looking while holding between the In-Sync and Out-of-Sync conditions, learning was only apparent in the In-Sync condition. This implies, for the Low LO Performing infants, that dynamic object views alone cannot promote learning and that they must coincide with the production of the label. Interestingly, these infants looked at the experimenter manipulation in the Look-Only condition substantially more (Look-Only: 65%; In-Sync and Out-of-Sync: 34-39%), but this did not impact their learning in that condition. This supports the recent finding that caregiver generated object view variability is not sufficient for learning as indicated by not being connected to later vocabulary growth (Slone et al., 2019). The difference between self- and experimenter generated views provides support that the manipulation is crucial while label occurs to promote learning.

These findings highlight the role of manipulation in an object-label mapping task for a subset of infants. The current study did not have any visual clutter and during the presentation phase objects were always viewed in isolation. Thus, visual isolation should have provided an optimal learning moment. However, this was not the case. For the High LO Performer infants, it could be that the isolation alone was beneficial to learning, and they would still benefit from manipulation in a cluttered scene task.

The present study demonstrates great variability within the sample. The grouped means for the interaction between LO Performance group and condition do not capture the full effect (Figure 10 shows the variability at the level of the individual). The In-Sync condition for the High LO Performers and both groups for the Out-of-Sync condition had an average proportion change below zero. However, there is considerable variability in performance in these conditions. Over half of the infants in each of the described conditions showed an increase of looking, while the other half per condition showed a sharp decline.

It should be noted that the Out-of-Sync condition has more variability than the In-Sync condition. For the Low LO Performing infants, this can be explained by these infants greatly benefiting from the three-way synchrony, however, for some, manipulation after the label production could still be enough to produce better learning conditions. Close to half of the infants from the High LO Performance group did worse in the Out-of-Sync condition—it is not clear as to why these infants performed worse, given that the majority of the High LO Performing infants mapped the In-Sync objects well. The Look-Only condition is similar in the presentation style of the Out-of-Sync condition; what varies is the dynamic object views and the manipulation. The only differences are either the manipulation or the shift in who is creating the dynamic objects views to explain the variability between the High LO performers in the Look-

Only and Out-of-Sync conditions. The current data cannot explain this variability, but it leaves room for further questioning into the individual learning variation due to task dynamics.

Environmental and Research Implications

The benefit of synchrony between looking, holding, and hearing in the Low LO Performer group implies that these infants will take advantage of hearing labels while actively playing in their environment. In infants' day-to-day lives, caregivers often use pointing and labeling to attract attention to objects that are out the infants' reach. This scenario would provide a synchronous moment for hearing the label and seeing the object; however, for some infants, this is not sufficient. The benefit of intermodal synchrony between holding, hearing, and seeing calls for more active learning environments, connecting play with language and learning. While the groups of infants did not differ in vocabulary sizes, they do fundamentally demonstrate different ways of learning object-label maps. Thus, as long as the infants' environment includes various types of learning experiences (such as visual and manual), then infants should still demonstrate equivocal vocabulary growth. It is possible we do not see a difference of MCDI between learning groups because of the age of the sample, which was at minimum 18-months-old. Manual experiences emerge later on in development, more so with the onset of walking, and the infants in our sample had at least a couple months of walking experience. It could be that the Low Look-Only infants had lower vocabulary scores earlier in development when synchrony between hearing, holding, and seeing was not as frequent as visual-only experiences. Therefore, by the time infants reached the age range for the present study, the differences may have already evened out with their walking experience. The current data set cannot answer this question, however. These results call for examining learning processes earlier in development when infants' motor skills are less stable to see if there is a relationship with their vocabulary scores.

Within the laboratory setting, these findings of some infants not performing well in visual only experiences call into question the traditional practice of using 2-D images to assess object label mapping, word learning, and other aspects of language development. For some infants, this type of task will produce learning; however, for others, there could be a significant gap in performance creating a large degree of variability in the sample. While studies using this paradigm have found interesting and valid results, considering an alternative explanation such as lack of physical interaction could explain sample variability because the present study demonstrates visual only mapping does not benefit some infants.

The current field examining the intersection of motor engagement and language skills often relates different task dynamics (motor skills, dynamic object views) to broader outcomes, typically vocabulary size as reported by the MCDI. This approach looks only at the outcome, not the process. While studying the outcome of different developmental factors can lead to a greater understanding of developmental trajectories, it does not answer questions regarding the *how* and *why* of the developmental process.

Limitations and Future Directions

This study has several limitations. First, infants demonstrated uneven baseline looking to the target and distractor objects (Figure 6). The lack of distributed baseline looking is likely due to the uneven presentation between the labeled and unlabeled objects. Unlabeled objects were presented only once during the presentation phase, while the labeled objects were presented four times each. The uneven balance of labeled and unlabeled object allowed for a reduced the number of trials; therefore not fatiguing the infants. However, the method of calculating the proportion increase from the baseline to the window of analysis should account for the potential novelty effect.

Research shows that infants tend to learn better with self-selection guided by interest. Self-selection of objects was not in the present study. Thus, interest in the task and stimuli could have a significant impact on the results. In Yu and Smith (2012), the infants selected the objects they interacted with (likely due to interest) and then they were freely labeled. As work from the play literature shows, learning with toddlers is better if the caregiver follows the interest of the learner (Dunham et al., 1993). The present task had objects preselected into each condition. Therefore, it did not follow the naturally accruing interest the infants had with the objects. The lack of self-directed interest could explain the variability in the data. However, this limitation also provides strength. Even though infants did not select the objects to learn, there was still a benefit of labeling synchrony with looking, holding, and hearing.

The lack of detailed data on the dynamic object views could be a limitation. The current data assumes that if infants are looking and holding an object, then they are experiencing a dynamic view. It could be that infants have not yet developed the preference for the horizontal view (James et al., 2014), or are experiencing it. Therefore the connection between dynamic object views and object-label mapping may not be captured in this current data set.

A limitation in studying object-label mapping could be that the task is measuring working memory. Research has shown that when testing word learning over a five-minute gap, to ensure that the test is relying on long term memory, infants do not perform well in the delay condition (Horst & Samuelson, 2008). These results suggest that the infants are fast mapping the labels to the objects, but not retaining the words over time. Therefore, it could be that the present task is testing the infant's working memory of the object-label mapping since there is not a delay period between the presentation trials and the test trials. For the present study, however, I did not aim to test long-term acquisition of novel words. Instead, we look at the

foundational aspect of novel word learning, which requires being able to map a label onto an object.

Lastly, walking and talking onsets were self-reported from the caregivers, not measured in the lab. However, by 18 months, infants have experienced in walking. The self-report data had a broad range (Walking: 8-16 months; Talking: 4-15 months). This variability suggests that the lacking relationship between motor skills and task performance in the present study should not be strictly interpreted, as the variability in the sample leads to questions of the self-report validity. Future research should have formal motor milestone assessments.

One of the questions that remain from this study is why infants differ in their learning strategies. We see that attention drops to a visual-only presentation for some learners, but not for others. In the present study, it is clear how the task dynamics are impacting subsequent learning, but does not explain the observed variations across the infants. Future studies need to address *why* infants benefit from different learning experiences. The present study captures the individual differences, but cannot explain them.

Conclusion

The present study shows an interesting effect on how infants respond to different task dynamics during a novel object-label mapping task. Some infants, our High LO Performer infants, do not need intermodal synchrony between hearing, holding, and seeing. For these infants, looking and hearing is sufficient for learning. It could be that the need for three-way synchrony is more critical for different developmental demands. For example, we know that infants and young children cannot always rely on three-way synchrony (i.e., for a dangerous object that infants should not manipulate, but do eventually learn the object label). However, when infants do struggle with visual only experiences, providing a synchronous moment between holding the object, looking at it, while hearing the label can significantly boost object

label mapping. These findings provide support for a dynamic systems view of object-label mapping, where multiple factors can interact to produce a learning environment that results in success in the task.

REFERENCES

- Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development*, 67(6), 3135–3153.
- Bergelson, E., & Aslin, R. N. (2017). Nature and origins of the lexicon in 6-mo-olds. *Proceedings of the National Academy of Sciences*, 114(49), 12916–12921. <https://doi.org/10.1073/pnas.1712966114>
- Bertenthal, B. I., Campos, J. J., & Barrett, K. C. (1984). Self-produced locomotion. In *Continuities and discontinuities in development* (pp. 175–210). Springer.
- Bruner, J. (1985). Child's talk: Learning to use language. *Child Language Teaching and Therapy*, 1(1), 111–114.
- Chomsky, N. (1967). Recent contributions to the theory of innate ideas. In *A Portrait of Twenty-five Years* (pp. 31–40). Springer.
- Chomsky, N. (2014). *Aspects of the Theory of Syntax* (Vol. 11). MIT press.
- Connell, J. P., DiMercurio, A., & Corbetta, D. (2017). Dynamic Systems Theory. In J. Vonk & P. D. T. Shackelford (Eds.), *Encyclopedia of Animal Cognition and Behavior* (pp. 1–8). https://doi.org/10.1007/978-3-319-47829-6_1594-1
- Corbetta, D., DiMercurio, A., Wiener, R. F., Connell, J. P., & Clark, M. (2018). How Perception and Action Fosters Exploration and Selection in Infant Skill Acquisition. In J. M. Plumert (Ed.), *Advances in Child Development and Behavior* (pp. 1–29). <https://doi.org/10.1016/bs.acdb.2018.04.001>
- Corbetta, D., & Snapp-Childs, W. (2009). Seeing and touching: The role of sensory-motor experience on the development of infant reaching. *Infant Behavior and Development*, 32(1), 44–58. <https://doi.org/10.1016/j.infbeh.2008.10.004>

- Corbetta, D., Williams, J. L., & Haynes, J. M. (2016). Bare fingers, but no obvious influence of “prickly” Velcro! In the absence of parents’ encouragement, it is not clear that “sticky mittens” provide an advantage to the process of learning to reach. *Infant Behavior and Development*, 42, 168–178.
- DiMercurio, A., Connell, J. P., Clark, M., & Corbetta, D. (2018). A Naturalistic Observation of Spontaneous Touches to the Body and Environment in the First 2 Months of Life. *Frontiers in Psychology*, 9, 2613. <https://doi.org/10.3389/fpsyg.2018.02613>
- Dunham, P. J., Dunham, F., & Curwin, A. (1993). Joint-attentional states and lexical acquisition at 18 months. *Developmental Psychology*, 29(5), 827.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... Stiles, J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development*, 59(5), i–185. <https://doi.org/10.2307/1166093>
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., & Reznick, J. S. (2014). *MacArthur-Bates Communicative Development Inventories: user’s guide and technical manual*. PB Brookes.
- Franchak, J. M., van der Zalm, D. J., & Adolph, K. E. (2010). Learning by doing: Action performance facilitates affordance perception. *Vision Research*, 50(24), 2758–2765.
- Frick, J. E., & Richards, J. E. (2001). Individual Differences in Infants’ Recognition of Briefly Presented Visual Stimuli. *Infancy*, 2(3), 331–352. https://doi.org/10.1207/S15327078IN0203_3
- Gershkoff-Stowe, L. (2001). The course of children’s naming errors in early word learning. *Journal of Cognition and Development*, 2(2), 131–155.
- Gesell, A., Thompson, H., & Amatruda, C. S. C. (1934). *Infant behavior: Its genesis and growth*.

- Ghisletta, P., & Spini, D. (2004). An Introduction to Generalized Estimating Equations and an Application to Assess Selectivity Effects in a Longitudinal Study on Very Old Individuals. *Journal of Educational and Behavioral Statistics*, 29(4), 421–437. <https://doi.org/10.3102/10769986029004421>
- Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology*, 49(3), 182–194.
- Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, 71(4), 878–894.
- Gogate, L. J., Bolzani, L. H., & Betancourt, E. A. (2006). Attention to Maternal Multimodal Naming by 6-to 8-Month-Old Infants and Learning of Word–Object Relations. *Infancy*, 9(3), 259–288.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD, US: Paul H Brookes Publishing.
- He, M., Walle, E. A., & Campos, J. J. (2015). A Cross-National Investigation of the Relationship Between Infant Walking and Language Development. *Infancy*, 20(3), 283–305. <https://doi.org/10.1111/infa.12071>
- Hoff, E. (2003). Causes and consequences of SES-related differences in parent-to-child speech. In *Monographs in Parenting Series. Socioeconomic status, parenting, and child development* (pp. 147–160). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Horst, J. S., & Samuelson, L. K. (2008). Fast Mapping but Poor Retention by 24-Month-Old Infants. *Infancy*, 13(2), 128–157. <https://doi.org/10.1080/15250000701795598>

- James, K. H., Jones, S. S., Smith, L. B., & Swain, S. N. (2014). Young Children's Self-Generated Object Views and Object Recognition. *Journal of Cognition and Development*, 15(3), 393–401. <https://doi.org/10.1080/15248372.2012.749481>
- Lewkowicz, D. J. (1996). Perception of auditory–visual temporal synchrony in human infants. *Journal of Experimental Psychology: Human Perception and Performance*, 22(5), 1094.
- Libertus, K., & Violi, D. A. (2016). Sit to talk: relation between motor skills and language development in infancy. *Frontiers in Psychology*, 7.
- Little, D. Y., & Sommer, F. T. (2011). Learning in embodied action-perception loops through exploration. *ArXiv Preprint ArXiv:1112.1125*.
- Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, 38(1–2, Serial No 149), 136–136. <https://doi.org/10.2307/1165788>
- Oakes, L. M. (2009). The “Humpty Dumpty problem” in the study of early cognitive development: Putting the infant back together again. *Perspectives on Psychological Science*, 4(4), 352–358.
- Oudgenoeg-Paz, O. (2014). *Walk this way, Talk this way. Motor skills, spatial exploration, and the development of spatial cognition and language*.
- Oudgenoeg-Paz, O., Leseman, P. P., & Volman, M. (2015). Exploration as a mediator of the relation between the attainment of motor milestones and the development of spatial cognition and spatial language. *Developmental Psychology*, 51(9), 1241.
- Pereira, A. F., James, K. H., Jones, S. S., & Smith, L. B. (2010). Early biases and developmental changes in self-generated object views. *Journal of Vision*, 10(11), 22–22. <https://doi.org/10.1167/10.11.22>

- Pereira, A. F., Smith, L. B., & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin & Review*, 21(1), 178–185.
- Peters, A. M. (1977). Language Learning Strategies: Does the Whole Equal the Sum of the Parts? *Language*, 53(3), 560–573. <https://doi.org/10.2307/413177>
- Quine, W. O. van. (1964). On Simple Theories of a Complex World. In *Form and Strategy in Science* (pp. 47–50). https://doi.org/10.1007/978-94-010-3603-0_6
- Quine, W. V. (1960). Carnap and logical truth. *Synthese*, 12(4), 350–374.
- Rochat, P., & Goubet, N. (1995). Development of sitting and reaching in 5-to 6-month-old infants. *Infant Behavior and Development*, 18(1), 53–68.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 1926–1928.
- Salvi, G., Montesano, L., Bernardino, A., & Santos-Victor, J. (2012). Language bootstrapping: Learning word meanings from perception–action association. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(3), 660–671.
- Samuelson, L. K., & McMurray, B. (2017). What does it take to learn a word? *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(1–2).
- Slone, L. K., Smith, L. B., & Yu, C. (2019). Self-generated Variability in Object Images Predicts Vocabulary Growth. *Developmental Science*, 0(ja), e12816. <https://doi.org/10.1111/desc.12816>
- Thelen, E. (1989). The (re) discovery of motor development: Learning new things from an old field. *Developmental Psychology*, 25(6), 946.

- Thelen, E., Corbetta, D., Kamm, K., Spencer, J. P., Schneider, K., & Zernicke, R. F. (1993). The transition to reaching: Mapping intention and intrinsic dynamics. *Child Development*, 64(4), 1058–1098.
- Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, 24(1), 1–34.
- Thelen, E., & Smith, L. B. (1998). Dynamic systems theories. *Handbook of Child Psychology*.
- Thurman, S. L., & Corbetta, D. (2017). Spatial Exploration and Changes in Infant-Mother Dyads around Transitions in Infant Locomotion. *Developmental Psychology*, 53(7), 1207–1221.
<https://doi.org/10.1037/dev0000328>
- Tomasello, M. (2010). *Origins of human communication*. MIT press.
- Walle, E. A., & Campos, J. J. (2014). Infant language development is related to the acquisition of walking. *Developmental Psychology*, 50(2), 336.
- Wang, M., Lekhal, R., Aarø, L., & Schjølberg, S. (2014). Co-occurring development of early childhood communication and motor skills: results from a population-based longitudinal study. *Child: Care, Health and Development*, 40(1), 77–84.
- Wiener, R. F. (2016). *The modulating role of motor action anticipation in relation to visual attention to a scene*.
- Williams, J. L., & Corbetta, D. (2016). Assessing the Impact of Movement Consequences on the Development of Early Reaching in Infancy. *Frontiers in Psychology*, 7.
- Williams, J. L., Corbetta, D., & Guan, Y. (2015). Learning to reach with “sticky” or “non-sticky” mittens: a tale of developmental trajectories. *Infant Behavior & Development*, 38, 82–96.
<https://doi.org/10.1016/j.infbeh.2015.01.001>

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, 125(2), 244–262.

APPENDICES

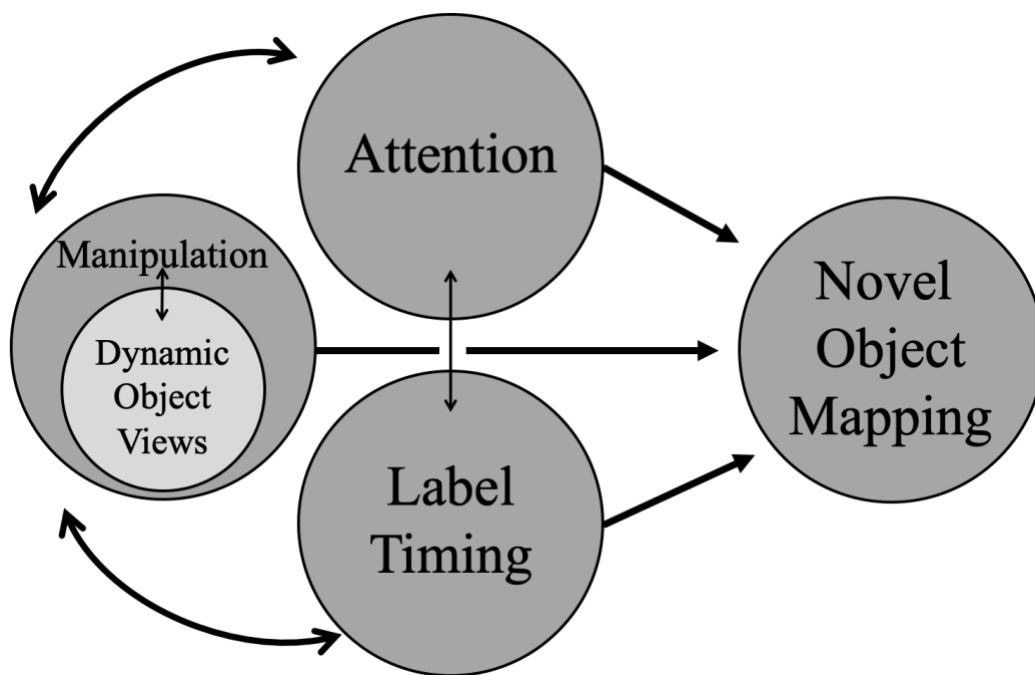


Figure 1. A model of novel object-label mapping using dynamic factors of a task. Manipulation should feed into and drive attention, as well as dynamic object views. All the factors together should feed into greater learning of object labels.

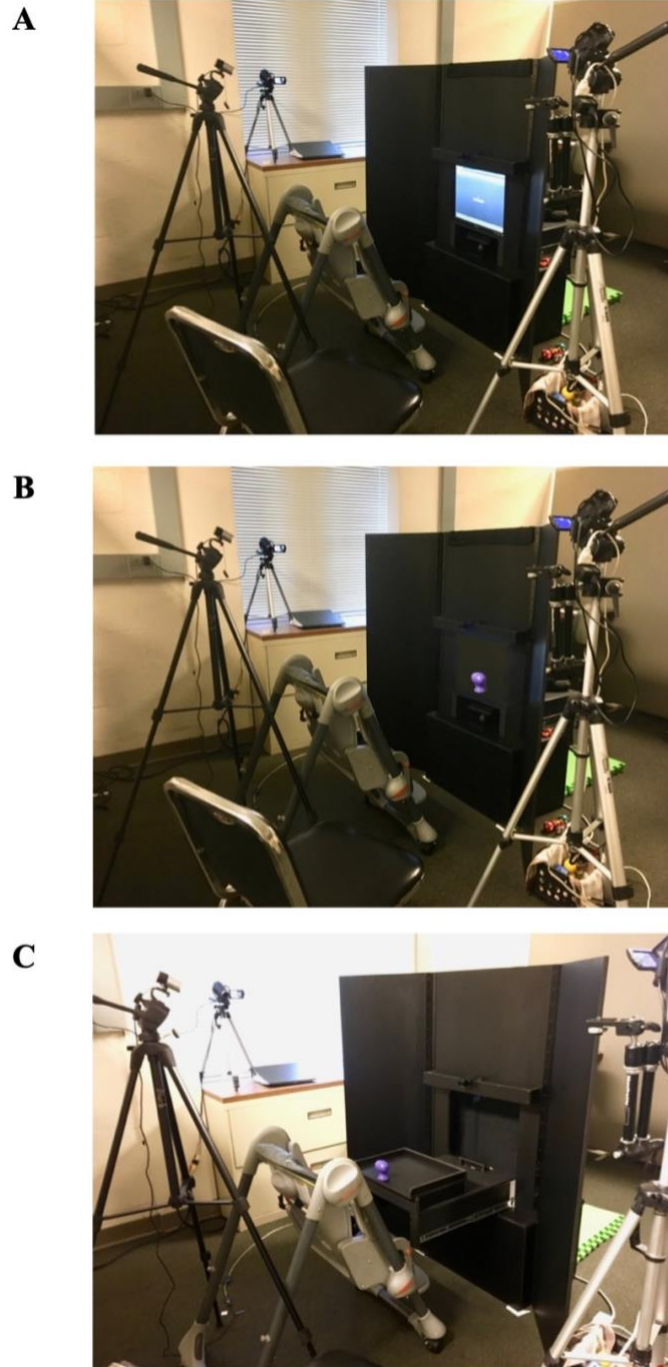


Figure 2. (A) The set-up for calibration when a computer screen is fitted into the window where presentation typically occurs. (B) All conditions received an out of reach presentation of the object in a single view. (C) For the In-Sync and Out-of-Sync conditions, the area for object presentation, flush with the theater during the presentation, is pushed out to be in the infants' reaching space for manipulation.

A



B



Figure 3. Novel stimuli used in the present study. (A) Labeled objects. (B) Unlabeled objects.

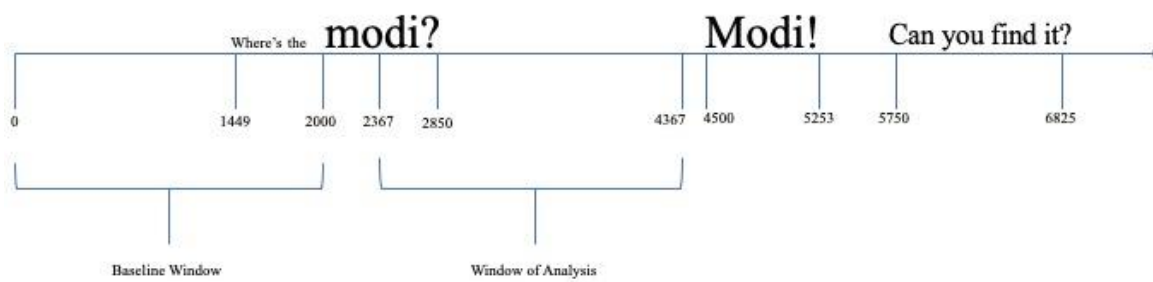


Figure 4. Timeline of auditory string.

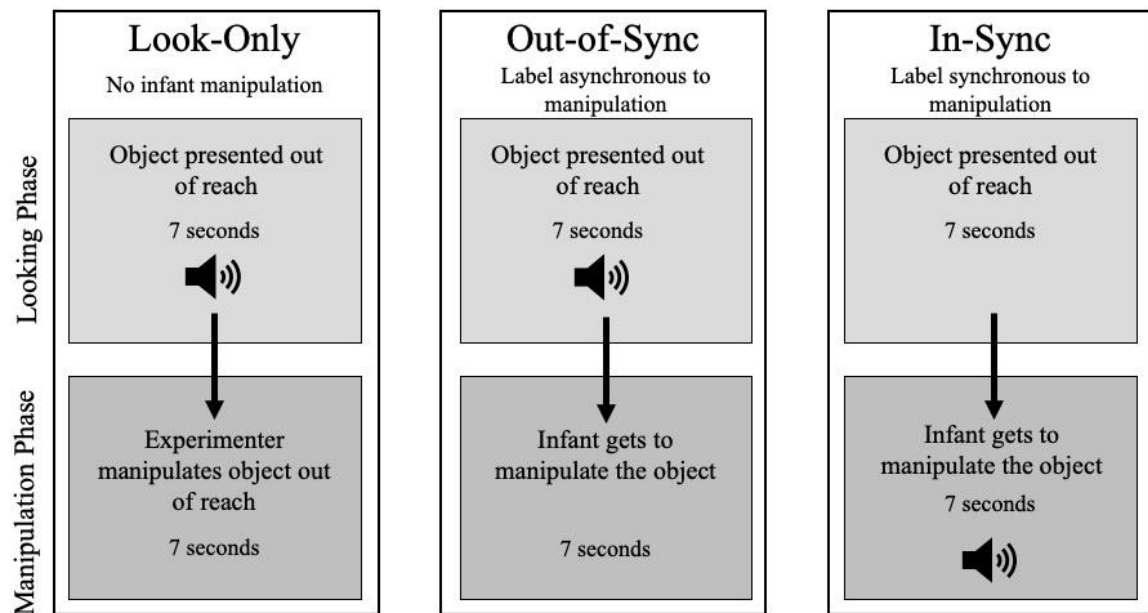


Figure 5. Description of the differences between the three within conditions. The speaker indicates when the auditory label was played.

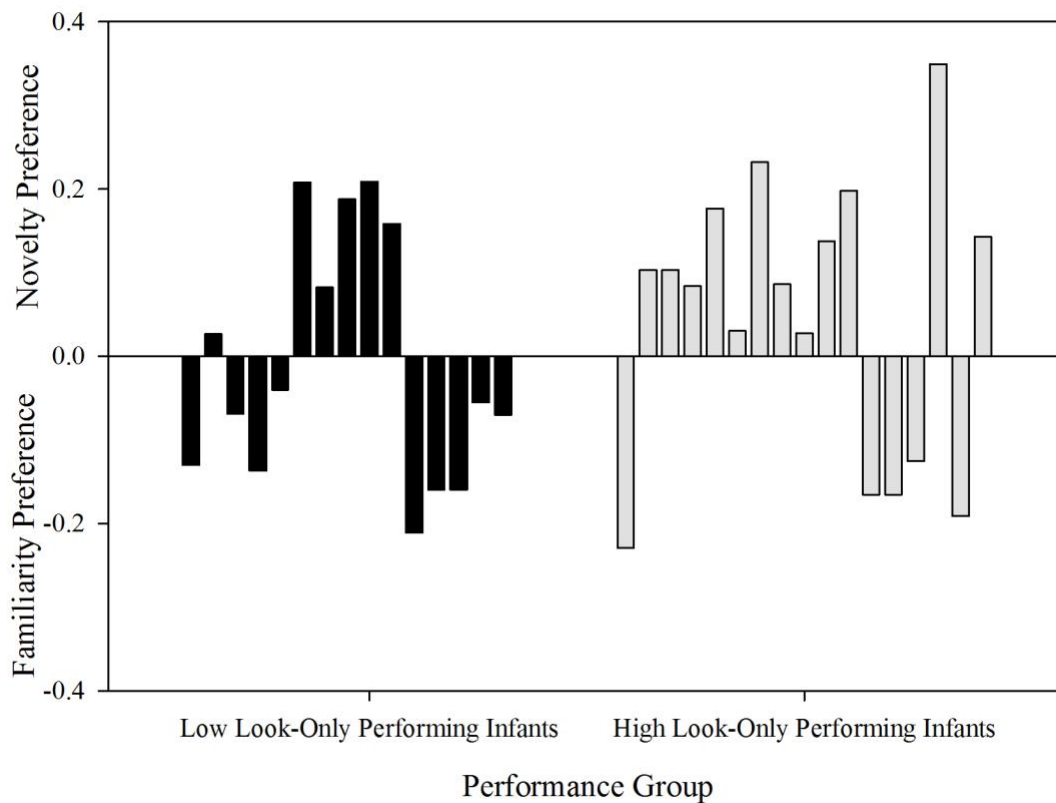


Figure 6. Looking preference during the baseline window of the test trial, 2000 ms. A positive value demonstrates a novelty preference while a negative value demonstrates a familiarity preference. A value of zero would represent even looking between the distractor and target object.

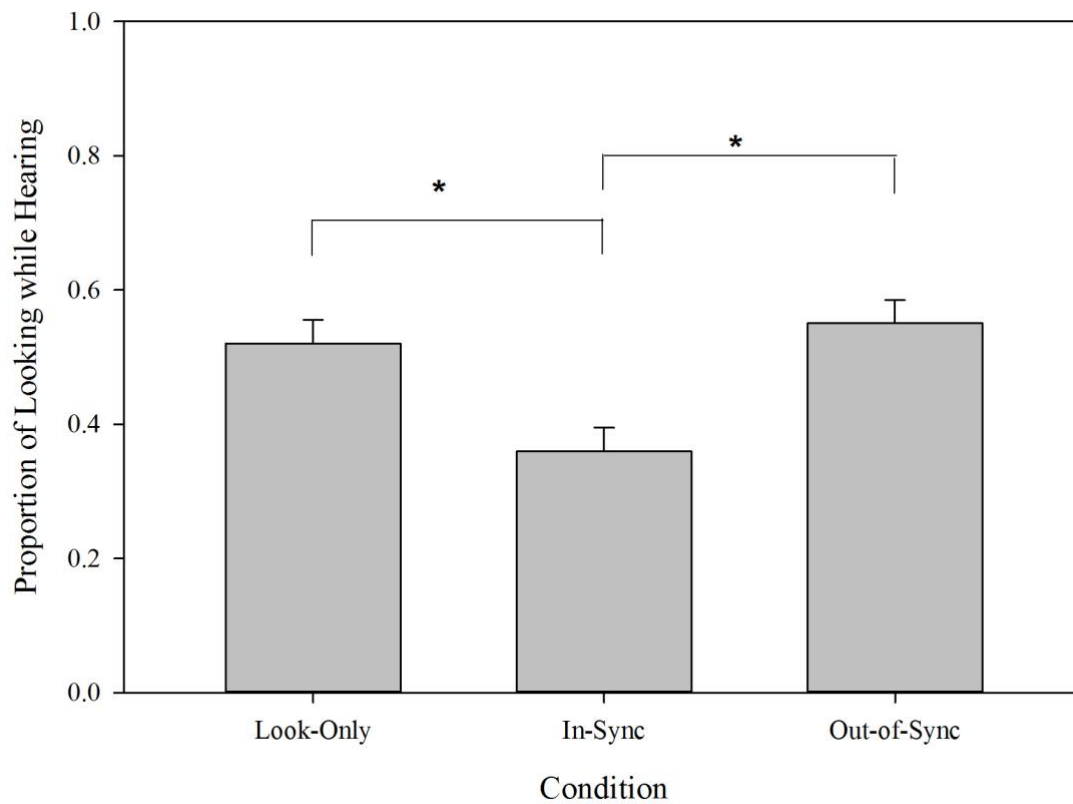


Figure 7. Main effect of condition for the proportion of looking while hearing from the presentation phase of the study.

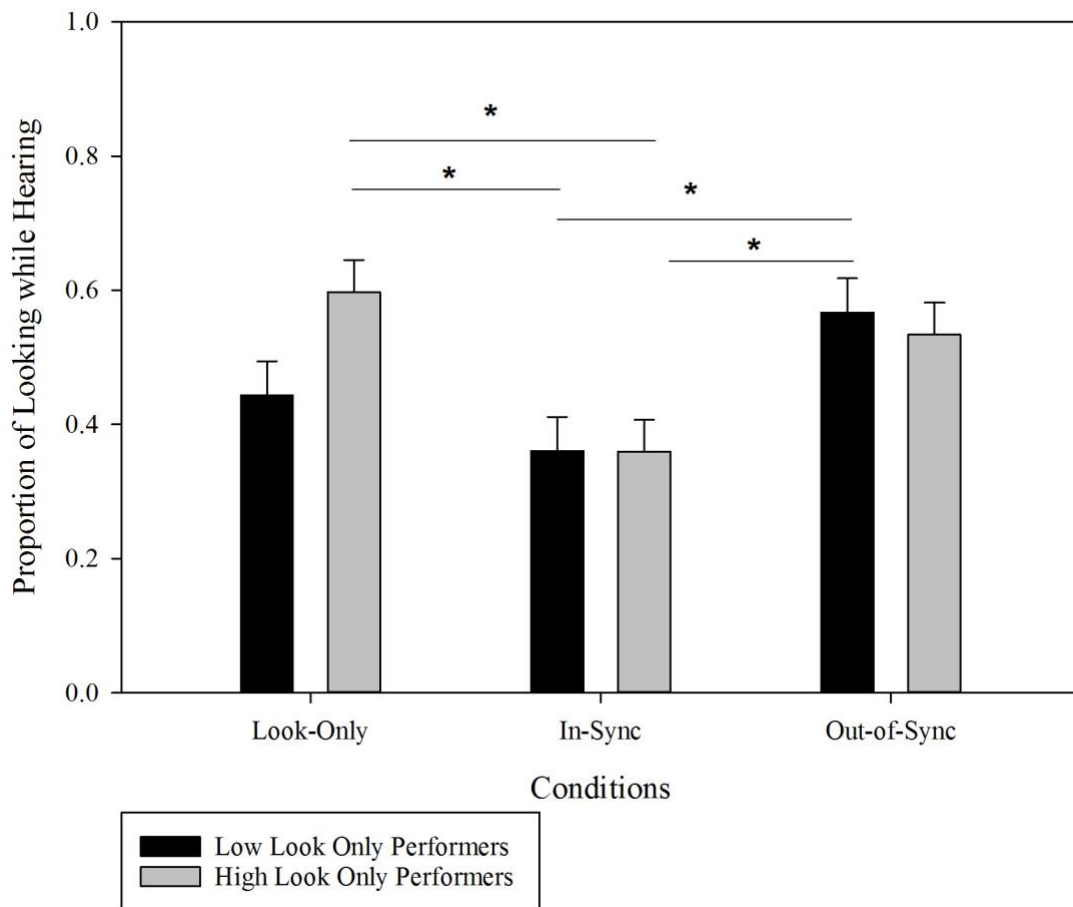


Figure 8. Proportion of looking while hearing during the presentation phase of the study by condition and between performance groups.

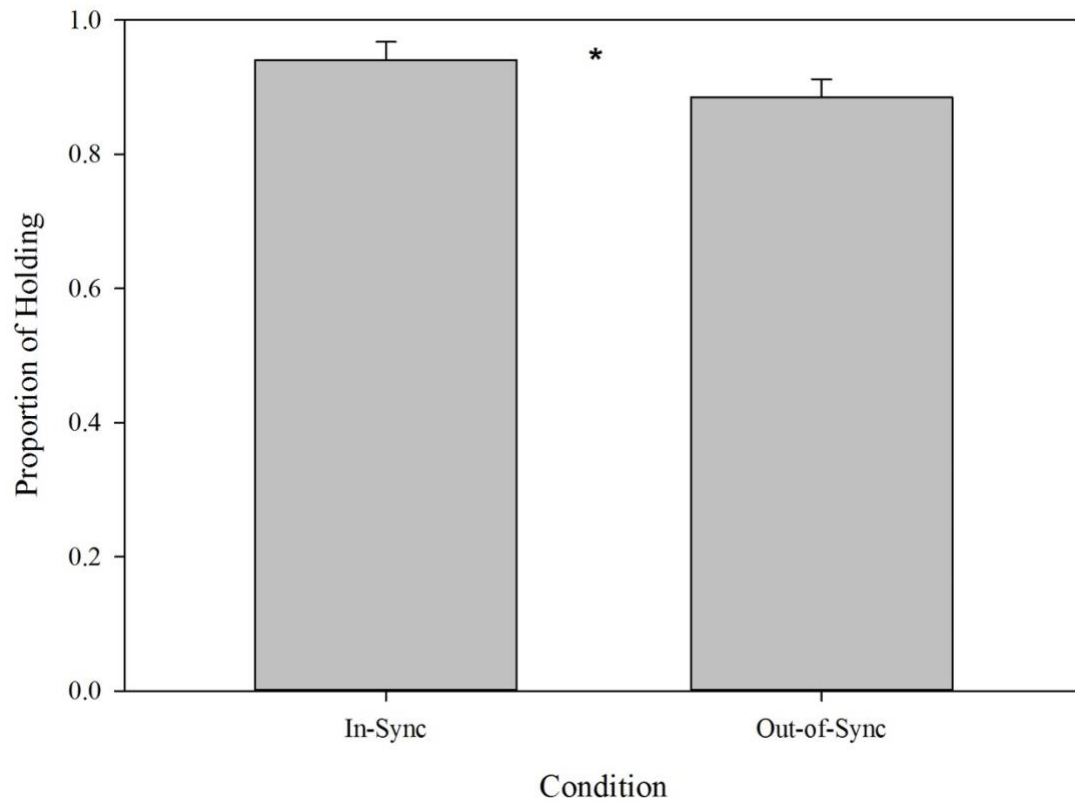


Figure 9. Proportion of time infants held the object during manipulation (In-Sync and Out-of-Sync) by condition during the presentation phase of the study.

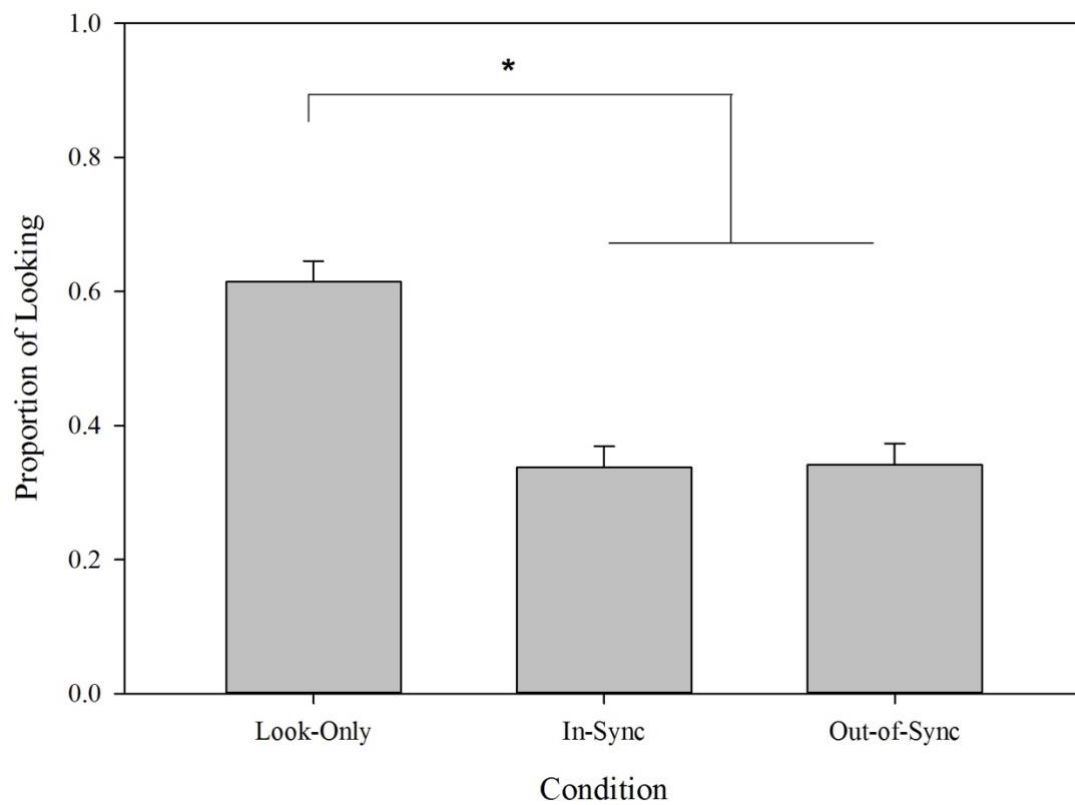


Figure 10. Proportion of looking towards the stimulus during self-manipulation (In-Sync and Out-of-Sync) or experimenter manipulation (Look Only) by condition during the presentation phase of the study.

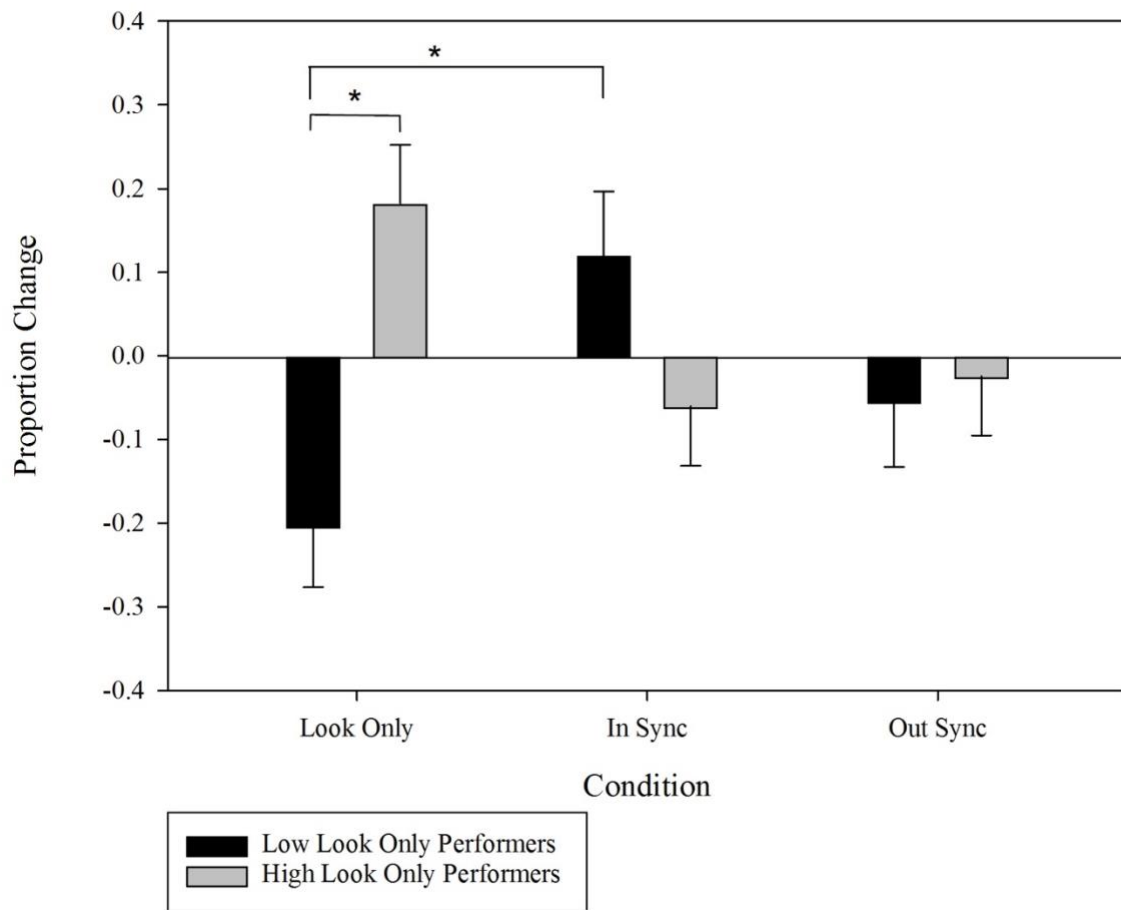


Figure 11. The proportion change of looking towards the target stimulus from the baseline portion of the test trial to the window of analysis by condition and between learning group. These values are based on the marginal mean estimates from the GEE.

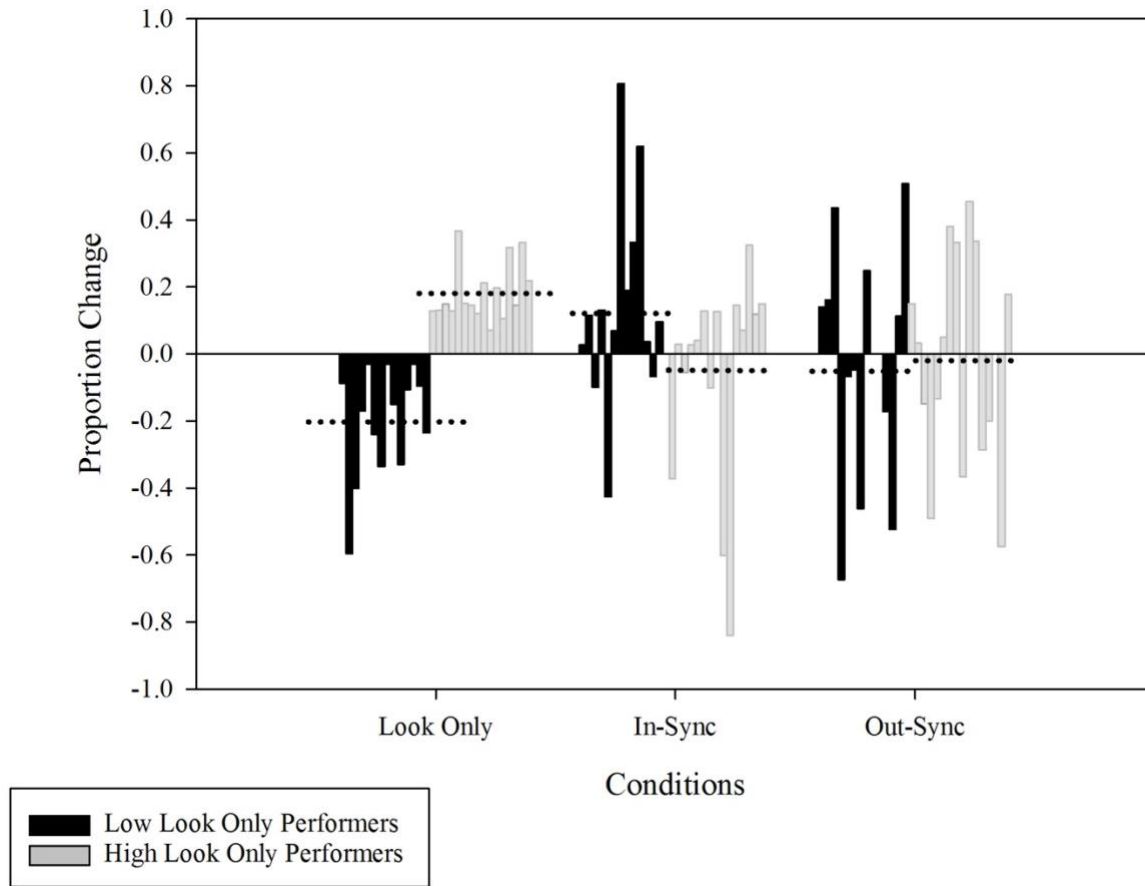


Figure 12. The proportion change from the baseline window of the test trial to the window of analysis at the level of the individual, between conditions. Each bar is a participant.

Table 1

Mean estimates between Performance groups and condition for presentation looking

Learning Groups	Condition	Mean	Std. Error
Low LO Performers	Look-Only	0.394	0.046
	In-Sync	0.460	0.046
	Out-Sync	0.482	0.045
High LO Performers	Look-Only	0.528	0.423
	In-Sync	0.494	0.423
	Out-Sync	0.499	0.423

Table 2

Mean estimates for looking and hearing during presentation trials

Condition	Mean	Std. Error
Look-Only	0.521	0.035
In-Sync	0.360	0.035
Out-Sync	0.550	0.035

Table 3

Mean estimates between Performance groups and condition for presentation looking and hearing

Learning Groups	Condition	Mean	Std. Error
Low LO Performers	Look-Only	0.443	0.051
	In-Sync	0.360	0.051
	Out-Sync	0.567	0.051
High LO Performers	Look-Only	0.598	0.477
	In-Sync	0.360	0.477
	Out-Sync	0.534	0.477

Table 4

Mean estimates for manipulation presentation trials

Condition	Mean	Std. Error
In-Sync	0.941	0.027
Out-Sync	0.885	0.027

Table 5

Mean estimates for looking during manipulation during presentation trials

Condition	Mean	Std. Error
Look-Only	0.615	0.031
In-Sync	0.338	0.031
Out-Sync	0.342	0.031

Table 6

Mean estimates between learning groups and condition

Learning Groups	Condition	Mean	Std. Error
Low LO Performers	Look-Only	-0.2026	0.0740
	In-Sync	0.1203	0.0762
	Out-Sync	-0.0542	0.0785
High LO Performers	Look-Only	0.1828	0.0692
	In-Sync	-0.0604	0.0710
	Out-Sync	-0.0242	0.07092

Table 7

Pairwise comparisons between the performance for group and condition

Group by Condition	Learner Group by Condition	Bonferroni Significance
Low LO Performer – Look-Only	In-Sync	0.044*
	Out-Sync	1.000
	High LO Performer – Look-Only	0.002*
	High LO Performer – In-Sync	1.000
	High LO Performer – Out-Sync	1.000
Low LO Performer – In-Sync	Look-Only	0.044*
	Out-Sync	1.000
	High LO Performer – Look-Only	1.000
	High LO Performer – In-Sync	1.000
	High LO Performer – Out-Sync	1.000
Low LO Performer – Out-Sync	Look-Only	1.000
	In-Sync	1.000
	High LO Performer – Look-Only	0.353
	High LO Performer – In-Sync	1.000
	High LO Performer – Out-Sync	1.000
High LO Performer – Look-Only	Low LO Performer – Look-Only	0.002*
	Low LO Performer – In-Sync	1.000
	Low LO Performer – Out-Sync	0.353
	In-Sync	0.248
	Out-Sync	0.866
High LO Performer – In-Sync	Low LO Performer – Look-Only	1.000
	Low LO Performer – In-Sync	1.000
	Low LO Performer – Out-Sync	1.000
	Look-Only	0.248
	Out-Sync	1.000
High LO Performer – Out-Sync	Low LO Performer – Look-Only	1.000
	Low LO Performer – In-Sync	1.000
	Low LO Performer – Out-Sync	1.000
	Look-Only	0.866
	In-Sync	1.000

VITA

Abigail DiMercurio was raised in Suwanee, Georgia by Sam and Debbie Dimercurio. She has one sister, Nicole DiMercurio. She attended the University of Georgia and received a Bachelors of Science degree in Psychology in 2016. Abigail is currently enrolled at the University of Tennessee in a doctoral program in Experimental Psychology.