Seeking new words: Active word learning in infants and children

By

Martin Zettersten

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Psychology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2020

Date of final oral examination: 4/22/2020

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

Martha Alibali, Professor, Psychology
Joe Austerweil, Assistant Professor, Psychology
Heather Kirkorian, Associate Professor, Human Development and Family Studies
Gary Lupyan, Associate Professor, Psychology
Jenny R. Saffran, Professor, Psychology

Table of Contents

Acknowledgements

An acknowledgements section feels like too small and low-dimensional a space to capture the immense gratitude I feel to all of the people who have helped me along my graduate school path. Know that if you do not find your name on these pages, I will find a way to express my gratitude to you in some other way.

I'd first like to thank all of the friends and loved ones who supported my initial, scary step to come to the United States for graduate school. Lauri, Steffi, Corinna, Amir, Tabby, Maddy, the Romans, bro and sis, and many others: I would not have been able to make that step without your help, support, and love.

Thanks to all of the people who made graduate school time in Madison so wonderful. Lab mates, Erica Wojcik, Viridiana Benitez, Ron Pomper – we're not a couple, but we were always in it together – , Molly Lewis, Christine Potter, Brianna McMillan, Tianlin Wang, Chiara Santolin, Desia Bacon, Haley Weaver, Sarah Oakley, Erin Long, Jing Shen, Rachel Reynders, Lynn Perry, Pierce Edmiston, Marcus Perlman, Christina Schonberg, Hettie Roebuck, Courtney Venker, Eileen Haebig, Anthony Goodwin, Tristan Mahr, Lilia Rissman, Ellise Suffill, and all of the other grad students, post-docs, research assistants and friends of the Infant Learning Lab and the Lupyan Lab: thanks for the incredible community you created, and for the countless things I learned from all of you. To Jason Samaha, Matt Cooper Borkenhagen, Regina Lapate, Clint Jensen, Mark Koranda, Drew Sheldon, Adrienne Wood, Rista Plate, Chris Cox, Elise Hopman, Steve Schwering, Katharine Scott, David Menendez, Tina Woodard, and many, many other buds and collaborators in Madison: you helped me learn how to work and live as a scientist, helped me grow, and made life unquantifiably more fun. To frisbee, to MuFa, to Ring and CUBE, to

Boomtown Pandas: thanks for keeping me sane and for giving me something more concrete and tangible than that strange abstract thing called a PhD to chase over the years.

I would like to highlight the intellectual contributions and support of many colleagues who helped make this work possible. Thanks in particular are due to my co-author on the project presented in chapter 3, Koeun Choi. Cory D Bonn, Chuck Kalish, Celeste Kidd, Terry Regier, Sam Wass, and Andrew Young all contributed valuable advice, resources, and ideas to projects presented here. I would like to thank my committee members: to Martha Alibali, Joe Austerweil, and Heather Kirkorian, for being consistent sources of support and insight over the years. I am grateful to Gary Lupyan, for many things, but in particular for encouraging me to see the big picture when developing scientific questions, and to never shy from a technical or methodological challenge. Finally, I would like to thank Jenny Saffran – I could not have asked for a better mentor, as a scientist and as a person. I will do my very best to pass on all of the many qualities I admire in your mentorship.

I would also like to extend my thanks to Heather Davis and the Madison Children's Museum for being so accommodating and helpful over the years and to all of the children and families who participated in the studies presented here.

My greatest gratitude is reserved for my Mom and Dad, for their unwavering support and belief in me throughout the years, and for giving me the freedom and confidence to become an active learner. To my Mom, among many things, thanks for teaching me to always stay positive and make the best of any situation (even if there's a global pandemic unfolding as you finish your dissertation); to my Dad, for teaching me that a good story is in the telling.

Finally, thanks to Danielle, for supplying me with cats, succulents, tea and a desk while writing my dissertation – and for everything else.

Abstract

Young word learners are faced with the immensely difficult task of determining the meaning of thousands of new words based on ambiguous input. How do children learn word meanings given limited and incomplete information? One way that children might confront this task is by intervening in their environment in ways that support learning. Throughout development, children are more than simply passive absorbers of information – they actively construct their own learning environment. This ability could have important consequences for how children simplify the problem of learning new word meanings. We present the results from four studies investigating how infants and children seek out new information during word learning. Chapter 2 presents evidence from children (3-8 years of age) and adults that learners will systematically seek to reduce ambiguity about novel object-label associations. In chapter 3, we find that children (3-5 years of age) are motivated to sample more informative object-label associations and make selections that are tuned to their past experience. In chapters 4 and 5, we present the results from two lines of work asking whether even infants (aged 17 to 21 months) systematically sample information about novel words. In order to study infants' sampling behavior during a novel word learning task, we developed novel gaze-contingent eye-tracking methods that allowed infants to trigger labeling events on a screen. Preliminary findings from these studies provide mixed or inconclusive evidence regarding infants' early information-seeking strategies. In general, we find evidence that children systematically select words that support gaining new information, and that the tendency to reduce ambiguity during word learning becomes more robust over development.

Keywords: active learning; word learning; language development; self-directed learning; sampling; cross-situational word learning

List of Figures and Tables

List of Figures and Tables (cont.)

Chapter 1: Children as active word learners

1. Active learning

1.1. Plato's touchscreen

Let's start with a dark, dystopian thought experiment. Consider two children whose only exposure to the world is through a screen.[1] Perhaps this could be something like Plato's cave, with representations of the screen appearing in front of the child on a cave wall (Plato, 1997).[2] The images that flit across the screen are the entirety of these children's input. While children are provided with rich visual input from the outside world, they have no ability to control or manipulate this visual experience: they are purely passive observers of the visual world. Now, imagine that instead of those images being simply passively absorbed, the child could actively control elements on the screen – the screen interactively responds to the child's gaze, attention, gestures, and motor movements. As images appear, children can focus on different distinct entities and interact with them however they choose. As they reach out and touch them, events on the screen respond contingently to the child's selections. When they select a dog-like image, the word "dog" plays from the screen. When they move their attention to a penguin image, the penguin responds by flapping its wings and the child hears the word "penguin". The world the child experiences comes alive in response to their actions. Intuitively, the child who is able to interactively engage with the visual input on screen in this manner would have far richer learning experiences.

---

[1] While there is a rich tradition within psychology and philosophy of mind of framing questions through thought experiments, developmental psychologists have developed a reputation as having perhaps some of the more ethically dubious flights of imagination (see e.g., https://www.smbc-comics.com/comic/2014-05-08). I consider this thought experiment as engaging with this intellectual heritage.

[2] Of course, Plato's allegory of the cave is an analogy devised to broach different questions - fundamental questions on human nature, epistemology, and the organization of political life, questions that far outstrip the ambitions of the questions raised here. I only wish to invoke Plato's cave here in the spirit of conjuring an image in the reader's mind.

While this thought experiment, thankfully, does not make direct contact with realistic learning scenarios (though see e.g., Held & Hein, 1963), it clarifies an important aspect of infants' and children's experience. Children actively engage with their environment from an early age and are invariably in an environment where that engagement will lead to informative experiences. From the causal contingencies of physical objects – pressing a button leads it to light up, hitting a ball makes it roll – to a variety of social and linguistic contingencies – reaching for your parent makes them pick you up, holding a toy leads your parent to name that toy –, acting on the world leads to a host of learning opportunities for children.

The idea that children's self-directed engagement with their environment has a central role to play in development has a long and varied tradition in cognitive developmental theory (Bruner, 1961; Piaget, 1955; Vygotsky, 1978). More recently, this general notion has been taken up in theoretical ideas arguing that curiosity (Gottlieb, Oudeyer, Lopes, & Baranes, 2013; Kidd & Hayden, 2015; Oudeyer & Smith, 2016) and hypothesis-testing (Gopnik & Wellman, 2012; Schulz & Bonawitz, 2007) drive development. One common underlying theme is that children are self-directed learners, in the sense that they can exert control over their learning input (Gureckis & Markant, 2012). One need not appeal to a version of Plato's cave, updated for the touchscreen era, to find infants actively manipulating their environment in ways somewhat analogous to the scenario described above. As a simple example, imagine an infant being read to by their parent. At a pause in the story, the infant points to the picture of an owl on the page, causing the parent to label the animal 'owl'. In this scenario, the infant has exerted control over their learning environment by eliciting information from their caregiver. Parent-child dynamics such as these, and information-seeking behaviors more generally, likely have a pervasive effect on the structure of children's learning environment (Smith, Jayaraman, Clerkin, & Yu, 2018).

The goal of this dissertation is to examine how children collect new information when encountering new words. If children are active word learners, what kind of active learners are they, and why?

1.2. What is active learning?

What makes learning "active"? In a recent book bringing together research on active learning from many different domains in development, Ganea and Saylor (2018) introduce active learning as goal-directed information-seeking:

> Active learning is characterized by a goal directed search for information. At its base, active learning involves the ability to identify gaps in one's knowledge, skills for seeking the missing information and the inclination to do so. (Saylor & Ganea, 2018: p. 4)

This definition has two aspects that I would like to slightly tweak for the purposes of this dissertation. First, requiring active learners to be able to identify knowledge gaps seems to set metacognitive reasoning skills as a prerequisite for active learning. I will not treat identifying gaps in knowledge as a prerequisite for active learning, instead embracing a broader sense of information-seeking[3] that aims to establish continuities between infant and child learning. However, the interpretation of many of the experiments in the following chapters will assume that infants have access to at least some metric that tracks their uncertainty about different stimuli, for example (see also chapter 6 for a broader discussion of the development of meta-cognitive reasoning and its relationship to active learning).

Second, active learning is usually thought to be a goal-directed process. A crucial part of this definition is that a child is not only seeking out new information, but is seeking out new

---

[3] I will use terms such as "information-seeking" and "information sampling" somewhat interchangeably throughout. The process of collecting new information can be usefully – and somewhat a-theoretically – conceptualized as a sampling from the environment (Fiedler, 2000; Fielder, 2008).

information in the service of achieving some goal (for instance, learning a new word or discovering a new property about a newly encountered animal). However, most, if not all, behaviors can be construed as goal-directed, given a broad enough interpretation of what the goal of some action might be. Even seemingly random behaviors can be construed as goal-directed – the goal here might be "generating novel or varying perceptual events". A key aspect in understanding different information-seeking strategies thus lies in characterizing the kinds of goals infants and children might pursue while actively seeking information. In the present work, our particular focus will lie on information-seeking in the service of learning. However, for present purposes, I will remain agnostic about whether and when learners are explicitly representing their goal as reducing uncertainty or improving learning when sampling new information.

2. Information-seeking in infants and children

2.1. Infant looking preferences and the Goldilocks effect

Much work in cognitive development relies on methods that show infants' discrimination of novel and familiar stimuli. The traditional view of infant looking times is that they are reactions to visual or auditory experience, which may be driven by exogenous factors (e.g., how salient a stimulus is) or endogenous factors (e.g., how robustly a stimulus is encoded in memory) (Aslin, 2014). More recently, infant looking behavior has begun to be re-conceptualized as a more active process (Kidd & Hayden, 2015; Kidd, Piantadosi, & Aslin, 2012). On this model, infants' looking behavior may reflect an active attempt to sample information from the environment (Carvalho, Vales, Fausey, & Smith, 2018).

A key result in understanding infants' gaze behavior as a more active process is the so-called Goldilocks effect (Kidd et al., 2012; Kidd, Piantadosi, & Aslin, 2014). Both in the visual and in the auditory domain, infants appear to prefer events that are 'just-right' in terms of their predictability: neither perfectly predictable nor completely unpredictable. Piaget (1952) hypothesized that "the subject looks neither at what is too familiar, because he is in a way surfeited with it, nor at what is too new, because this does not correspond to anything in his [schemes]" (p. 68). This idea connects with a broad finding from the curiosity literature showing that infants and adults demonstrate a U-shaped pattern of effort in seeking to explore novel, complex stimuli events and stimuli (Berlyne, 1966; Brennan, Ames, & Moore, 1966; Kidd & Hayden, 2015). For instance, in Kidd et al. (2012), infants viewed objects disappearing and reappearing behind a screen. By varying how predictable the pattern of reappearance of an object was from behind a particular screen, Kidd et al. obtained a measure of a given event's predictability or complexity. As an example, if an object can appear from behind one of two screens, an extremely predictable event is one in which an object appears from behind screen 1 after having appeared repeatedly from screen 1 on previous events (e.g., creating the sequence 1–1–1–**1**). On the other end of the continuum, if an object suddenly appears from behind screen 2 after having only appeared from behind screen 1 (i.e., the sequence 1–1–1–**2**), then the event is much more surprising. An event can also lie in between these two extremes, creating a pattern that has some variability, but is also somewhat predictable (e.g., 1–2–1–**2**). Crucially, in Kidd et al. (2012), the predictability or complexity of a particular event within a pattern influenced how long infants would continue to watch the event sequence. Infants showed a U-curve preference, with infants looking longest to patterns that were neither too predictable nor too unpredictable (i.e., events such as 1-2-1-**2** in the example above). This U-shaped curve held for every individual infant, not just for the group of participants overall (Piantadosi, Kidd, & Aslin, 2014).

Why might infants – and learners more generally – look longer to events that have intermediate complexity? In conceptualizing infants' looking as an active sampling process, curiosity researchers suggest that this effect results from infants being drawn to reduce uncertainty about upcoming events (Henderson, 2017; Itti & Baldi, 2006; Kidd & Hayden, 2015). Infants can easily form predictions about low complexity events, and therefore have little need to reduce uncertainty about what will happen next. It is also difficult to reduce uncertainty about high complexity events because upcoming events remain unpredictable. Events with intermediate complexity, on the other hand, may allow for the greatest amount of uncertainty reduction. However, the findings from infant looking time studies such as these allow for alternative explanations that do not appeal to uncertainty. For instance, it is possible that the attentional pattern is explained by some other low-level principle, such as a more general arousal principle, that leads infants to maintain a constant level of intermediate arousal. What other evidence is there that infants and children actively seek information to reduce uncertainty about the world around them?

2.2. Seeking new information to reduce uncertainty and ambiguity

Compelling evidence for information-seeking in infants comes from studies that ask how infants respond to surprising events. In one set of studies (Stahl & Feigenson, 2015), 11-month-old infants were shown events that either conformed to their expectations about the world or did not. For instance, an object was pushed off of a small cliff and either fell to the ground (knowledge-consistent event) or appeared to remain hovering in the air (knowledge-inconsistent event). Infants subsequently mapped a novel property to an object better if the object was involved in a knowledge-inconsistent event than if the object was involved in a knowledge-consistent event. Most interesting were the patterns of behavior infants showed in interacting with the objects after knowledge-consistent or knowledge-inconsistent events. When infants were given the chance to explore the objects they had just seen, they not only spent more time investigating objects involved in knowledge-inconsistent events - they also generated new information that could help explain the events they had just experienced. For instance, if an object had previously hovered in the air rather than falling to the ground, infants were more likely to repeatedly drop the object, apparently testing the previously observed hovering property. This study demonstrates that infants even at a young age actively seek out new information that can aid in explaining surprising events.

As children grow older, recent research suggests that they will actively structure their play in a way that is well-suited to gaining new information and reducing uncertainty (Cook, Goodman, & Schulz, 2011; Denison, Bonawitz, Gopnik, & Griffiths, 2013; Schulz, Gopnik, & Glymour, 2007; van Schijndel, Visser, van Bers, & Raijmakers, 2014). Two factors that can lead children to seek new information are experiencing *inconsistent* evidence and *ambiguous* or *confounded* evidence.

Research on causal learning suggests that preschool children are motivated to explain and explore novel events that are inconsistent with past experience (Legare, 2012; Legare, Gelman, & Wellman, 2010). In one study (Legare, 2012), children were exposed to “blicket detectors”, simple boxes that lit up whenever a novel object (labeled a “blicket”) was placed on it. Children were subsequently exposed to events in which objects previously observed to be “blickets” and “non-blickets” both did not trigger the box to light up. Children were motivated to explore and generate hypotheses about the “blicket” object that failed to light up the box, i.e. the event that was inconsistent with past experience in the experiment, rather than other events that were consistent with past evidence.

Preschool children are also motivated to investigate events that provide ambiguous evidence (Cook et al., 2011; Schulz & Bonawitz, 2007). For instance, in one study (Schulz & Bonawitz, 2007), a child and an experimenter first played with a toy with two levers that made two different puppets pop up. In one condition, the confounded condition, the two levers were always pressed together, causing both puppets to pop up simultaneously, so that children received no information about what each lever's role was in the puppets appearing. In other conditions, the causal role of the levers was de-confounded such that children could infer that each lever caused a different puppet to appear. After this initial play phase, a novel toy was revealed, and children were left alone to play with either the old toy or the new one. Children in the confounded condition preferred to continue playing with the old toy, while children in conditions in which there was no uncertainty about the causal structure of the old toy chose to spend more time playing with the newly introduced toy.

Together these findings suggest that preschool children and even infants spontaneously seek more information when there is higher uncertainty regarding the underlying causal structure of previously experienced events. Children are active information- and explanation-seekers, often targeting information that is helpful in reducing uncertainty about inconsistent or ambiguous past events.

3. Seeking new information about words

3.1. Children as active participants in word-learning processes

The notion that children are active participants in their word learning has a rich history in theories of language development. One prominent advocate for the importance of infants' intrinsic curiosity and motivations in word learning is Lois Bloom (L. Bloom, 2000a, 2000b; L. Bloom, Margulis, Tinker, & Fujita, 1996; L. Bloom, Tinker, & Scholnick, 2001). Bloom memorably argued that children's motivations and intentions play a crucial role in how the word learning events are constructed in a commentary in an edited volume collecting different theories of early word learning:

> I cannot help but note that I am the only one represented in this volume who listens to children learning to talk—who actively watches children's spontaneous acts of expression and interpretation as opposed to observing their responses to manipulations of word-learning events in an experiment. Children's spontaneous behaviors have by and large been ignored or, even worse, dismissed. But, I suggest, with all due respect, that we ignore children's spontaneous behaviors—what they *do* in the everyday events of their lives—at the peril of the theories that we construct to explain those behaviors. What children do in their activities of daily living has everything to do with learning words. **Children do not just wait around for *other* people to construct the word-learning scenario for them. Not on your life. Instead, *they create the word-learning process themselves*. The words they learn are words *they* want to learn, the words *they* need to learn. They are the words that are relevant to what *they* have in mind.** (Bloom, 2000b: p. 165; emphasis in bold mine)

This view of the children's early word learning foregrounds children's active participation in the word learning process. If children are active word learners, what does past research tell us about when and how children seek information about new words?

Past research has documented how endogenous attentional factors shape how infants and toddlers encode and track new word meanings. For example, one line of work has documented how object salience (roughly defined as the degree to which an object captures children's attention in a given context) and object novelty affect children's word learning. Children will typically map novel words to the most salient object in a given context (Samuelson & Smith, 1998). Similarly, children tend to map new words to novel objects (Markman & Wachtel, 1988) and are biased towards choosing novel objects when selecting the referent of novel words (Horst, Samuelson, Kucker, & McMurray, 2011). A classic result is that when children are presented with a new label in the context of a familiar object (e.g., a cup) and a novel object, children (and adults) will tend to map the new word to the new object (Halberda, 2006; Lewis, Cristiano, Lake, Kwan, & Frank, 2020), a phenomenon typically referred to as mutual exclusivity (Markman & Wachtel, 1988). Salience and novelty also interact in complex ways during children's word learning. In one study, the degree to which children successfully mapped a new word to a novel object in a mutual exclusivity task was shown to depend on the salience of the familiar competitor items (Pomper & Saffran, 2019). When children experienced a new word and object together with a highly salient, familiar object (such as a cat), they were less successful at associating the novel word with the novel object than if the novel object occurred together with a familiar object with low salience (such as a box), in part because they were slower to attend to the novel object in the presence of an attention-capturing item. The degree to which children are motivated to attend to different objects in their environment, e.g. due to their salience or novelty, has consequences for how they learn new words.

3.2. Pointing as a window into early information-seeking about words

While many studies have investigated how factors that shift children's attention to different objects shape learning, less is known about what influences whether and how children seek information about novel objects and their labels. One prominent avenue for approaching this question is through children's early pointing behavior. Children's gesture has an important role in the development of language (Goldin-Meadow, 2007; Goldin-Meadow & Alibali, 2013), with some researchers describing pointing as the "royal road" to language (Butterworth, 2003). Early pointing behavior is a powerful predictor of later language ability: infants' gesture behavior at 14 months is a strong predictor of later vocabulary size at 3 and 4 years of age, over and above other predictors such as children's and parent's word use at 14 months (Rowe & Goldin-Meadow, 2009; Rowe, Özçalişkan, & Goldin-Meadow, 2008). One reason why children's pointing relates to language learning is that pointing creates particularly informative learning instances. Caregivers often provide contingent, follow-in labeling responses in reaction to infants' pointing (Gros-Louis & Wu, 2012; Wu & Gros-Louis, 2014). Thus, pointing can be construed as a potentially powerful tool for infants to elicit labeling information.

Growing evidence suggests that pointing serves an epistemic or interrogative function in infancy, allowing infants to strategically seek new information about objects and their labels (Begus & Southgate, 2012; Kovács, Tauzin, Téglás, Gergely, & Csibra, 2014). For instance, when 16-month-old infants interact with an experimenter who correctly labels familiar objects and expresses certainty in labeling objects, they are more likely to point to novel objects than when interacting with an uncertain, less knowledgeable experimenter (Begus & Southgate, 2012). By 18 months, infants' pointing behavior appears to be – at least in some contexts - targeted towards obtaining label information about novel objects: in one study, 18-month-olds were less likely to continue pointing to a novel object after hearing a label, while infants' reaching behavior was not differentially affected by whether or not objects were labeled (Lucca & Wilbourn, 2018b).

Infants' pointing also has consequences for how well they learn new information about the objects they direct their behavior toward. Infants at 18 months learn novel labels for objects better after pointing towards them than after referencing them in other ways, such as reaching towards them (Lucca & Wilbourn, 2018a). This comports well with results from other studies demonstrating that 16-month-old infants remembered actions performed on a novel object better if they learned about an object they pointed towards during a selection phase than if they learned about an object they had not selected (Begus, Gliga, & Southgate, 2014). Together, these results suggest that between the ages of one and two, infants are motivated to hear labels for novel objects, often using pointing to specifically elicit labeling information, and that their information-seeking behavior supports the encoding of the labels and functions of novel objects.

3.3. Early question-asking behavior

Children's information-seeking and curiosity becomes more salient as they grow older and begin to ask increasingly targeted questions about the world around them, in particular about the objects and categories they encounter (Chouinard, 2007; Jiminez, Sun, & Saylor, 2018; Legare, Mills, Souza, Plummer, & Yasskin, 2013; Mills, Legare, Grant, & Landrum, 2011; Nelson, Divjak, Gudmundsdottir, Martignon, & Meder, 2014; Ruggeri & Lombrozo, 2015). Research based on corpus and diary studies estimates that preschool children ask an average of 77 information-seeking questions per hour when actively engaged with an adult (Chouinard, 2007), though the tendency to ask information-seeking questions is likely to vary considerably between individuals (Jiminez et al., 2018; Kurkul & Corriveau, 2018). One notable fact about children's early question-asking is that they frequently formulate questions requesting labels (e.g., "What's that?") or questions attempting to constrain the meaning of words (e.g., "What's a jack-o-lantern?"). These types of queries are the most frequent type of questions that children ask between the age of 18 months (around 60% of children's questions) and 3 years of age (around 20% of children's questions) as estimated from the CHILDES corpus, and continue to account for 12%-24% of children's questions as they approach the age of 5 (Chouinard, 2007; Jiminez et al., 2018). As children grow older and gain the ability to ask explicit questions about the world around them, they appear to use this ability to actively learn about new words.

3.4. Implications of active information-seeking strategies for early word learning

While these studies document that infants and children actively seek out information about new words, there remain fundamental questions about how children go about sampling new word meanings, and what motivates children's information-seeking in general (Coenen, Nelson, & Gureckis, 2019). What kinds of sampling patterns do infants and children show when they actively seek out information about words? Why do children seek out new information about words? And how do these sampling behaviors relate to later learning? Answering these fundamental questions about what type of information infants and children are motivated to seek has the potential to speak to more general questions about the relationship between children's early information-seeking behavior and later language development.

These questions are particularly important when considering that children must learn the meanings of words despite inherent ambiguity in any word learning situation (Quine, 1960). Children often encounter new words in complex social interactions surrounded by novel objects and events that offer many potential meanings and referents for novel words (Clerkin, Hart, Rehg, Yu, & Smith, 2017; Medina, Snedeker, Trueswell, & Gleitman, 2011; Smith et al., 2018). While children are well-equipped with statistical learning skills (Smith & Yu, 2008; Suanda, Mugwanya, & Namy, 2014) and social-cognitive abilities (Frank, Goodman, & Tenenbaum, 2009; Frank, Tenenbaum, & Fernald, 2013; Shafto, Goodman, & Frank, 2012; Tomasello & Akhtar, 1995; Tomasello, Carpenter, Call, Behne, & Moll, 2005; Yurovsky & Frank, 2017) to aid in confronting the complexity of the word learning task, children's ability to actively intervene on and structure their word learning environment may have a central role to play in explaining why children are such successful word learners (Smith et al., 2018).

Computational analyses suggest that self-directed sampling can vastly simplify the task of learning new words and categories under uncertainty, so long as learners employ strategies that increase the frequency of learning about word meanings that have higher uncertainty or are more ambiguous (Hidaka, Torii, & Kachergis, 2017; Oudeyer, Kachergis, & Schueller, 2019; Settles, 2012). For example, in one computational model of fast-mapping object-label associations (Hidaka, Torii, & Kachergis, 2017), successfully learning an adult-sized vocabulary (around 60,000 words) required an unrealistic number of sampling trials when learning events were assumed to be drawn randomly from a Zipfian word distribution. However, when the model implemented an active learner who preferentially selected less frequently encountered object-label associations, learning was sped up by several orders of magnitude. Other computational models have implemented active word learning in terms of curiosity mechanisms that preferentially sample objects that are expected to lead to the largest increases in accuracy (Twomey & Westermann, 2017). Sampling learning events based on this active selection mechanism leads to large increases in word learning speed and accuracy compared to randomly selected learning events, e.g. when learning to map labels to objects in visual scenes (Keijser, Gelderloos, & Alishahi, 2019). An intriguing result from these simulations is that active sampling mechanisms not only increase eventual accuracy, but also lead to more consistent or robust outcomes across simulations compared to simulations where learning events are selected randomly. These analyses suggest that at least in principle, active sampling mechanisms that make selections based on potential learning gain or uncertainty reduction significantly speed up and simplify the problem of tracking object-label associations for the learner.

Being an "active" word learner thus has the potential to substantially ease the complex task of learning word meanings under uncertainty. The central question then becomes – to what degree do infants and young children actively seek new information about words, and what types of sampling strategies do they exhibit when learning new words under uncertainty? Do children preferentially construct more informative word learning events, in a similar manner to the selection mechanisms implemented in computational models of active word learning?

4. The present work

The present work asks how infants and children sample information about words and what consequences their sampling strategies have for learning. In chapter 2, we present a series of experiments testing whether children and adults will attempt to disambiguate word meanings when presented with ambiguous word learning situations. If learners encounter confounded or ambiguous evidence about which object a novel word refers to, a particularly powerful learning strategy would be to construct learning events that disambiguate confounding evidence. Across four experiments, we test whether both adult and child learners (3 – 8 years of age) are motivated to sample disambiguating evidence when learning novel object-label associations. Next, chapter 3 investigates the effects of children's ability to actively control their word learning experience on subsequent learning by comparing children's learning performance in an "active" condition in which they choose which words to learn to "passive" conditions in which participants cannot control their learning experience. The guiding questions are whether children sample more informative object-label associations, and whether their sampling strategies promote better word learning. In Chapters 4 and 5, we ask whether even infants systematically seek information as they encounter new words. We present studies with infants between 17 and 21 months of age investigating whether infants are sensitive to the potential informativeness of different word learning events. In both studies, infants' experience with novel object-label associations is first manipulated in a training phase, such that hearing one set of object-label associations becomes potentially more informative than hearing the other. In the studies in Chapter 4, we manipulate the past frequency of infants' experience with different object-label associations, while in Chapter 5, we manipulate the consistency with which infants have experienced particular object-label associations. Both studies employ novel gaze-contingent methods that allow infants to control their subsequent learning experience by triggering different events on the screen. Together, these studies present initial forays into understanding how infants and children sample word learning events when they can exert control over their learning curriculum, and what their sampling behavior might mean for how new words are learned.

Chapter 2: Sampling to reduce ambiguity during cross-situational word learning

Why do we seek out new information during learning? One proposal is that information-seeking behavior is driven by uncertainty reduction (e.g., Kidd & Hayden, 2015). A variety of studies have demonstrated that – at least in some contexts - children may be motivated to gather information to reduce uncertainty after ambiguous or surprising events (Schulz & Bonawitz, 2007; Stahl & Feigenson, 2015). For instance, infants and children may preferentially seek out information from social partners when they are more uncertain (Goupil, Romand-Monnier, & Kouider, 2016) or when confronted with ambiguous or incomplete information (Bazhydai, Westermann, & Parise, 2020; Hembacher & Frank, 2017; Vaish, Demir, & Baldwin, 2011). Understanding the nature of children's information-seeking strategies may provide key insights concerning how children are able to rapidly solve complex learning problems across development (Gopnik et al., 2017; Oudeyer & Smith, 2016).

A classic problem in word learning is how learners disambiguate the meaning of words in potentially ambiguous situations (Quine, 1960). One solution is that children can disambiguate word meanings by tracking co-occurrences of object-label pairs across multiple ambiguous situations (Smith & Yu, 2008; Suanda et al., 2014; C. Yu & Smith, 2007). This proposal would be particularly powerful when combined with selective information-seeking: if learners are motivated to sample object-label associations that remained ambiguous over the course of past learning, this may substantially improve word learning (Hidaka et al., 2017; Keijser et al., 2019). Past studies with children suggest that they are sensitive to referentially ambiguous situations, preferentially seeking information from social partners when confronted with referential ambiguity (Hembacher & Frank, 2017; Vaish et al., 2011) and that children may learn words better for objects they are more curious about (Ackermann, Hepach, & Mani, 2019; Lucca &

Wilbourn, 2018b). Among adult learners, active selection of label-object pairs during cross-situational word learning increases participants' accuracy compared to a passive condition in which random sets of objects are presented (Kachergis, Yu, & Shiffrin, 2013). However, we still know little about what sampling strategies adult and child learners display when given the opportunity to control their own learning input.

In the current work, we investigated whether adult and child learners actively seek information that aids in reducing ambiguity about the meaning of novel words. We manipulated the ambiguity of novel word mappings by varying the degree to which object-label pairs co-occurred with one another during cross-situational word learning (Experiments 1A, 1B, 2A) or whether children could use mutual exclusivity to disambiguate the referents of novel words (Experiment 2B). The central question was whether adults and children would choose to learn more about those items that reduce uncertainty about novel object-label associations.

Experiments 1A & 1B

Experiments 1A-B were designed to determine whether adult learners would seek information that aided in disambiguating reference. Participants completed a cross-situational learning task in which their goal was to learn a set of object-label associations by determining the referent of each label across training. Participants were then given the opportunity to select which object-label association they would hear on each subsequent learning trial. The central question was whether adults would make selections that reduced referential ambiguity. We collected data in an online experiment (Experiment 1A) and in an in-lab experiment (Experiment 1B) with similar designs. Ambiguity was manipulated to varying degrees across two conditions. In Experiment 1A, all participants were assigned to a Fully Ambiguous condition, in which one set of items remained ambiguous throughout the training phase, while object-label associations were disambiguated for the remaining items. In Experiment 1B, participants were randomly assigned to the Fully Ambiguous condition or the Partially Ambiguous condition, in which one set of items was manipulated to remain moderately ambiguous across training, while the remaining items were disambiguated as in the Fully Ambiguous condition. We predicted that participants would be motivated to select items that were manipulated to be more ambiguous in both conditions, with a stronger preference for ambiguous items in the Fully Ambiguous condition.

Method

Participants

For Experiment 1A, we recruited 31 participants through Amazon Mechanical Turk (8 female; mean age: 31.4 years, *SD* = 7.25; all native speakers of English). Three participants were excluded for not passing an initial auditory attention check (2) or for restarting the experiment (1). All participants were assigned to the Fully Ambiguous Condition (n = 28) and paid $0.75 for completing the study.

For Experiment 1B, 62 University of Wisconsin-Madison undergraduates (27 female; mean age: 19.1 years, *SD* = 1.01; 56 native speakers of English) participated for course credit and were randomly assigned to the Fully Ambiguous Condition (n = 28) or the Partially Ambiguous Condition (n = 34).

Stimuli

The object stimuli were 8 images of novel 'alien' creatures used in previous word learning studies (Partridge, Mcgovern, Yung, & Kidd, 2015). 8 novel word stimuli (*beppo*, *finna*, *guffi*, *kita*, *noopy*, *manu*, *sibu*, *tesser*) were recorded by a female native speaker of English and normalized in duration and average loudness. The association between each label and its target referent and the roles of the stimuli within a condition were randomized across participants. The stimuli were presented using a web-based experiment created using jsPsych (de Leeuw, 2015). A demo of each experiment can be accessed through a web browser (Experiment 1A: http://sapir.psych.wisc.edu/~zettersten/demos/Crossact/version1/crossact_v1.html; Experiment 1B: http://sapir.psych.wisc.edu/~zettersten/demos/Crossact/version2Demo/crossact_v2.html).

Design & Procedure

The experiment was split into a *Training Phase*, a *Sampling Phase*, *Test Phase*, and a *Production Test Phase*. In Experiment 1A, participants completed each of these four phases in order. In Experiment 1B, participants completed the Test Phase twice, once after the Training Phase and once after the Sampling Phase (i.e., the order was Training-Test-Sampling-Test-Production Test; see below for the rationale motivating this design).

Training Phase. Participants completed 24 cross-situational learning trials (2 blocks of 12 trials), presented in random order (Figure 2.1). Participants were instructed that their goal was to learn the association between eight novel labels and their referents. On each training trial, participants were presented with two referents and two labels. The labels appeared sequentially in random order, both visually and auditorily. Consequently, the association between a particular label and its referent remained ambiguous on any single trial, but could be disambiguated by aggregating information across trials. Each object and its label occurred 6 times across the 24 training trials.

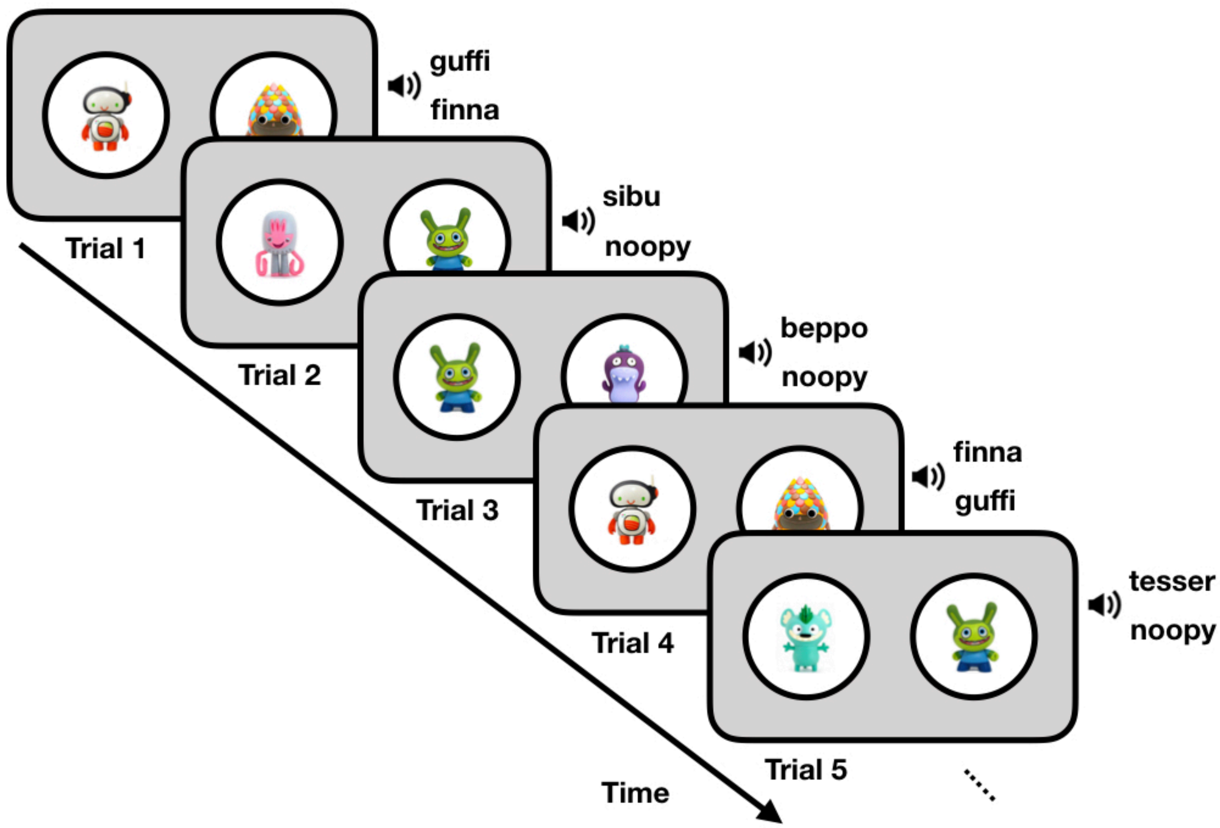

Figure 2.1. Overview over the training procedure.

We manipulated whether the object-label associations became disambiguated across trials during training, and therefore, how uncertain participants were at the onset of the *Sampling Phase* about the specific object-label pairs. In Experiment 1A, all participants completed the Fully Ambiguous condition. In Experiment 1B, participants were randomly assigned to one of two conditions: the <u>Fully Ambiguous condition</u> or the <u>Partially Ambiguous condition</u>. In the Fully Ambiguous condition, half of the object-label pairs remained ambiguous: two sets of two items were yoked together such that they were never disambiguated across training (<u>ambiguous</u> items; Figure 2.2, top left). The remaining items in the Fully Ambiguous condition were disambiguated across trials, occurring with three different object-label pairs (<u>disambiguated</u> items; Figure 2.2, right panel). In the Partially Ambiguous condition, two sets of two objects were grouped such that two specific objects co-occurred on 4 out of their 6 occurrences, but each occurred with one other object from the ambiguous object set on the remaining 2 trials (partially ambiguous items; Figure 2.2, bottom left). The other four objects were disambiguated as in the Fully Ambiguous condition. Note that across both conditions, participants saw each individual object and label equally frequently.

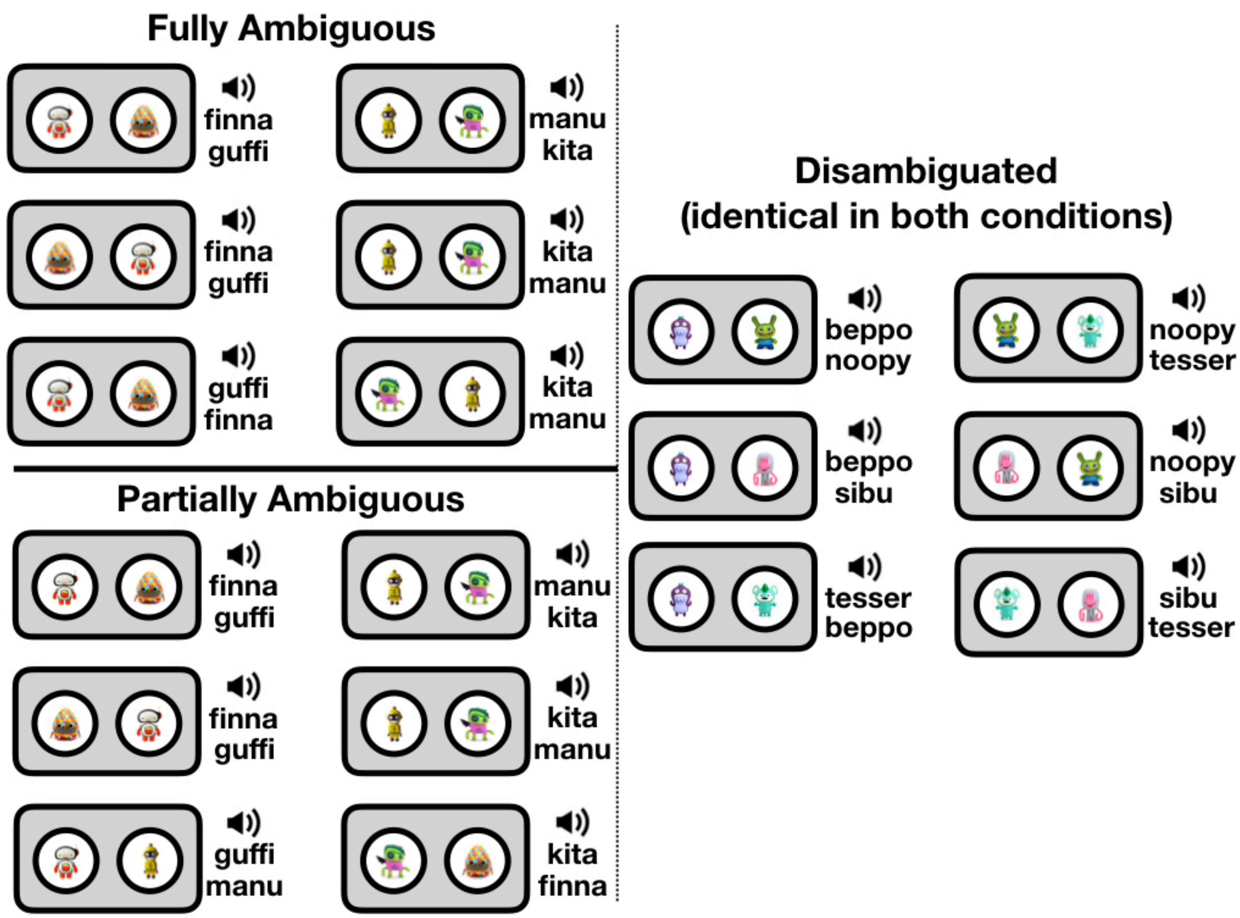

Figure 2.2. Overview over one block of the Training Phase for the Fully Ambiguous Condition and the Partially Ambiguous Condition.

Sampling Phase. Participants next completed four sampling trials. On each trial, all 8 objects appeared in randomized locations. Participants were instructed to select which of the 8 items they wanted to hear in the next cross-situational learning trial. After participants' selection, a second object was chosen at random from the remaining objects. The two objects and their labels then appeared together in a cross-situational word learning trial with the same structure as in the training phase.

Test Phase. Participants' knowledge of the object-label associations was probed in an 8-alternative forced choice (8-AFC) recognition test. On each test trial, all 8 objects appeared in randomized locations on the screen, along with one of the 8 labels. Participants were then asked to select the object that went with the label. No feedback was provided after a choice. Participants were tested on each label in random order, for a total of 8 recognition test trials.

In Experiment 1A, the Test Phase was presented once, following the Sampling Phase. In Experiment 1B, the 8-AFC test phase was presented twice: once immediately following the Training Phase (Test Block 1) and once immediately following the Sampling Phase (Test Block 2), as in Experiment 1A. Our rationale in testing participants' object-label knowledge twice was to further assess the consequences of participants' sampling behavior for learning. By testing participants' object-label knowledge before and after the Sampling Phase, we hoped to investigate whether participants' choices led to gains in test accuracy. While adding a test phase immediately following the Training Phase provided participants with additional exposure to the labels and the objects prior to the Sampling Phase, we reasoned that the additional test phase (Test Block 1) should not help participants learn the particular object-label associations, since all eight objects were presented with each label (i.e., no cross-situational learning could occur during the test phase) and participants did not receive feedback on their test choices.

Production Test Phase. After completing the final test phase, participants were asked to name each of the 8 objects from training. We included this additional test of participants' object-label knowledge in order to ensure we captured variability in participants' learning (e.g., in case participants performed at ceiling on the 8-AFC recognition test). On each trial, the image of one of the 8 objects appeared on the screen, and participants were asked to type the name for the object into the text box. The 8 objects were presented in random order for each participant.

Predictions

We predicted that participants would be more likely to choose to learn about the ambiguous items than about the disambiguated items in the sampling phase. For the Partially Ambiguous condition, we expected participants to have a weaker preference for ambiguous items over the disambiguated items, since adults accurately tracking the co-occurrence evidence could successfully learn all word-referent pairs. We did not predict large differences in test accuracy between items. One possible outcome was that test accuracy would be higher for items that were disambiguated during training. However, another possibility was that ambiguous items could be learned at comparable levels to disambiguated items if participants preferentially sampled ambiguous items.

Results

Sampling choices

Experiment 1A. We fit a logistic mixed-effects model testing participants' likelihood of making an ambiguous selection against a chance level of 0.5 using the lme4 package (version 1.1-21) in R (Bates & Maechler, 2009; R Development Core Team, 2019), including by-participant and by-participant random intercepts. Participants were more likely to choose ambiguous items than disambiguated items, b = .62, Wald 95% CI = [0.06, 1.17], z = 2.16, p = .03 (Figure 2.3). Participants chose an object from the ambiguous set on 62.5% (95% CI = [50.8%, 74.2%]) of trials. To test the robustness of this result, we also tested subjects' average proportion of ambiguous selections against the chance level of 0.5 in a non-parametric statistical test. A Wilcoxon signed-rank test found that the distribution of ambiguous selections significantly diverged from chance, $V = 152$, $p = .02$.

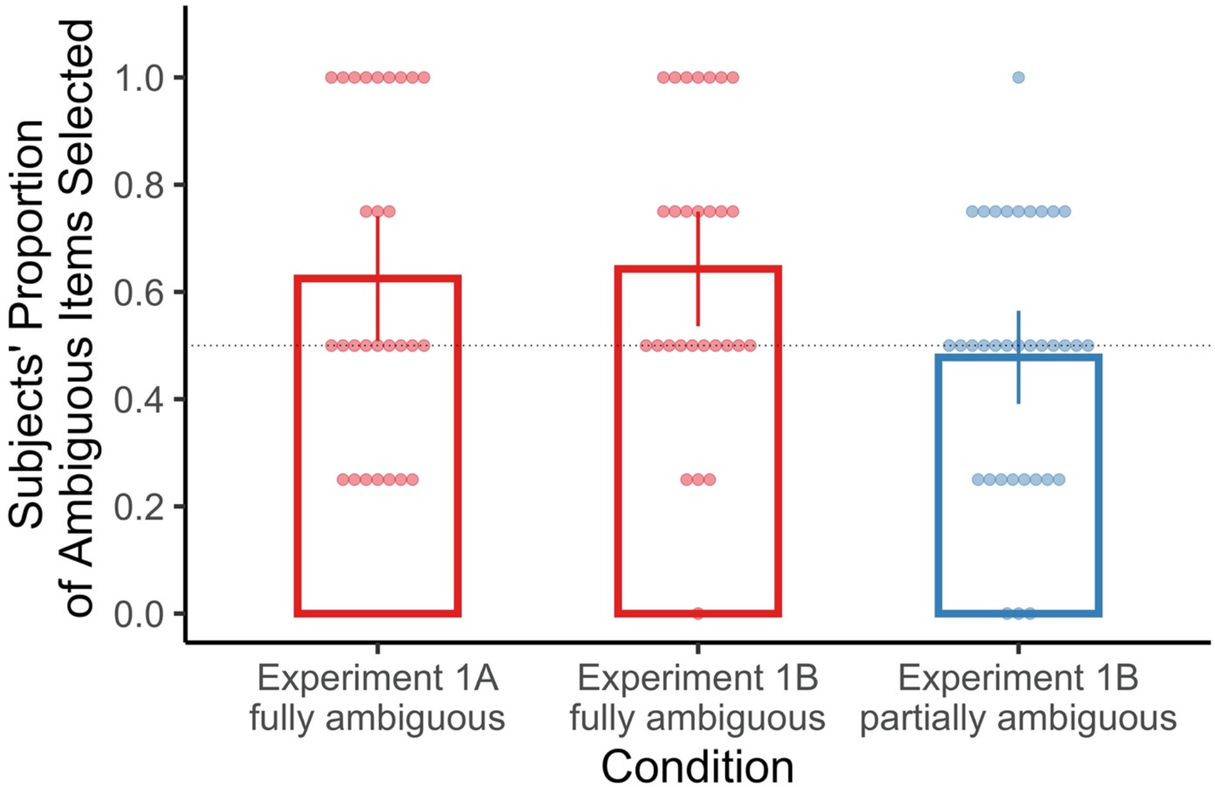

Figure 2.3. Proportion of more ambiguous items selected in each condition. Error bars represent 95% CIs.

Experiment 1B. We fit a logistic mixed-effects model predicting participants' likelihood of making an ambiguous selection from condition (centered: fully ambiguous = 0.5; partially ambiguous = -0.5) (Bates & Maechler, 2009; R Development Core Team, 2019), including by-participant and by-participant random intercepts. To test whether participants' likelihood of selecting an ambiguous item differed from chance within each condition, we refit the model while recoding the condition variable with each condition coded as zero (such that the intercept represents a test of the likelihood of an ambiguous selection against a chance level of 0.5 for the condition coded as zero, since logit(0.5) = 0).

Participants were significantly more likely to select a (fully or partially) ambiguous item in the Fully Ambiguous condition (M = 64.3%, 95% CI = [53.6%, 75.0%]) than in the Partially Ambiguous condition (M = 47.8%, 95% CI = [39.1%, 56.5%]), $b = 0.70$, 95% Wald CI = [0.14, 1.25], $z = 2.47$, $p = .01$. A Wilcoxon rank-sum test comparing subjects' proportion of ambiguous selections between the two conditions yielded similar results ($W = 632$, $p = .02$). Participants in the Fully Ambiguous condition selected ambiguous items more frequently than disambiguated items, replicating the result from Experiment 1A (b = 0.60, 95% CI = [0.19, 1.02], z = 2.85, p = .004; Wilcoxon signed-ranked test: $V = 140$, $p = .015$). Contrary to our prediction, participants in the Partially Ambiguous condition did not select (partially) ambiguous over disambiguated items (b = -0.09, 95% CI = [-0.45, 0.27], z = -0.50, p = .62; Wilcoxon signed-rank test: $V = 101$, $p = .59$).

8-AFC Test performance

Experiment 1A. Overall, participants successfully learned the object-label pairs, accurately selecting the correct referent in the Fully Ambiguous condition (M = 65.6%, 95% CI = [52.7%, 78.6%], chance = 12.5%). To compare accuracy for items that remained ambiguous during training to accuracy for items disambiguated during training, we fit a logistic mixed-effects model predicting trial-by-trial accuracy from item type (centered; ambiguous = 0.5; disambiguated = -0.5), including by-participant and by-item random intercepts and a by-participant random slope for item type. Accuracy for ambiguous items (M = 63.4%, 95% CI = [56.2%, 70.5%]; corrected within-participants; Morey, 2008) was marginally lower than accuracy for items that were disambiguated during training (M = 67.9%, 95% CI = [60.7%, 75.0%]), $z = -1.65$, $p = .10$ (see Figure 2.4A).

Experiment 1B. Participants demonstrated successful word learning in the 8-AFC recognition test that immediately followed the Training Phase (Fully Ambiguous condition: M = 65.2%, 95% CI = [55.6%, 74.7%]; Partially Ambiguous condition: M = 76.8%, 95% CI = [66.7%, 87.0%]) and in the (identical) test phase following the Sampling Phase (Fully Ambiguous condition: M = 72.8%, 95% CI = [61.6%, 84.0%]; Partially Ambiguous condition: M = 77.6%, 95% CI = [67.1%, 88.0%]). To investigate the relationship between Test Half, Item Type (ambiguous vs. disambiguated), and Condition, we fit a logistic mixed-effects model predicting trial-by-trial accuracy from the three-way interaction between these three predictors (centered), including all lower-order effects. We included by-participant and by-item random intercepts, and by-participant random slopes for Test Block, Item Type, and their interaction. There was a significant effect of Item Type (b = -1.33, Wald 95% CI = [-2.10, -0.56], z = -3.39, p < .001), indicating that participants generally performed better on items disambiguated during training across conditions, and a significant effect of Test Block (b = 0.82, Wald 95% CI = [0.11, 1.53], z = 2.27, p = .02), indicating that participants performed better on the second test block than the first. This effect of Test Block appeared to be driven mainly by a significant increase in accuracy for ambiguous items in the Fully Ambiguous condition from Test Block 1 M = 46.4%, 95% CI = [33.6%, 59.3%]) to Test Block 2 (M = 59.8%, 95% CI = [45.3%, 74.4%]), b = 1.03, Wald 95% CI = [0.18, 1.89], z = 2.38, p = .017 (see Figure 2.4B). There was also a significant interaction between Item Type and Condition, indicating that the difference in accuracy between disambiguated and ambiguous items was greater in the Fully Ambiguous condition than in the Partially Ambiguous condition (averaging across test block), b = -2.75, Wald 95% CI = [-3.87, -1.64], z = -4.83, p < .001 (see Figures 2.4B and C). All other effects, including the three-way interaction, were non-significant, all ps > .25.

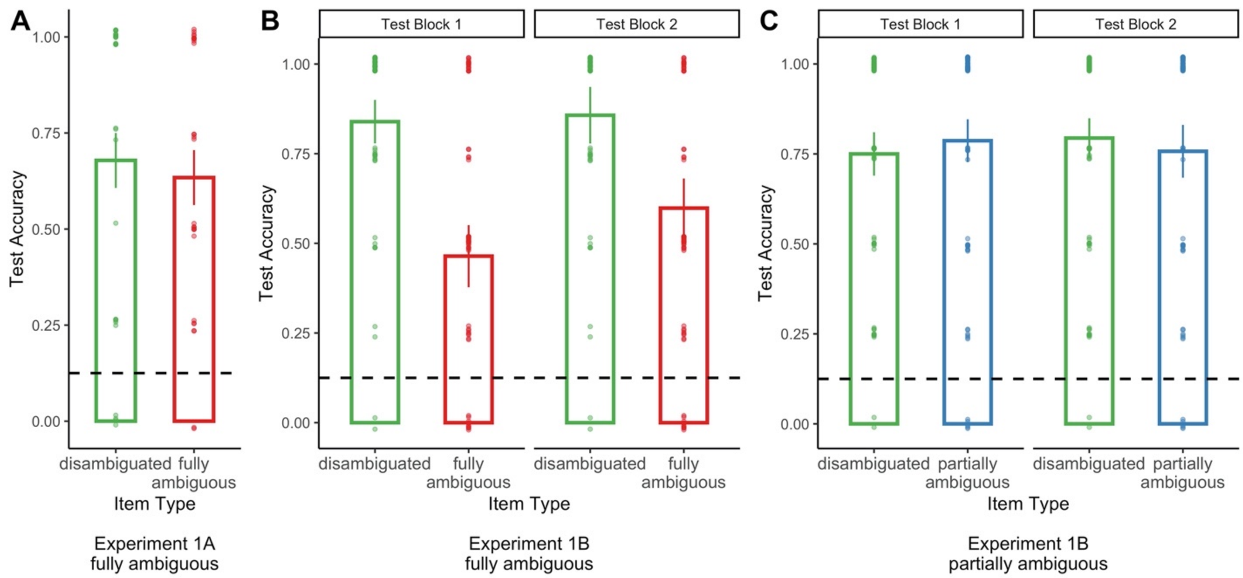

Figure 2.4. Test accuracy by item for (A) Experiment 1A, (B) the Fully Ambiguous condition in Experiment 1B, and (C) the Partially Ambiguous condition in Experiment 1B. The dashed line represents chance-level performance. Error bars represent within-participant 95% CIs (Morey, 2008).

Production Test Performance

Experiment 1A. Overall, participants produced the correct label for roughly half of the eight object-label pairs (M = 47.8%, 95% CI = [32.6%, 63.0%]). We fit the same logistic mixed-effects model as for the 8-AFC recognition test to compare performance on ambiguous and disambiguated test items. Participants did not differ in their accuracy at producing labels for ambiguous (M = 47.3%, 95% CI = [40.5%, 54.2%]) and disambiguated items (M = 48.2%, 95% CI = [41.4%, 55.1%]), $z = 0.07$, $p = .95$ (Figure 2.5A).

Experiment 1B. Participants produced the correct label for approximately half of the eight object-label pairs in both the Fully Ambiguous (M = 52.2%, 95% CI = [39.0%, 65.4%]) and in the Partially Ambiguous condition (M = 55.9%, 95% CI = [44.6%, 67.2%]). To investigate whether performance differed across Condition and Item Type, we fit a logistic mixed-effects model predicting whether participants produced the correct label from Condition (centered), Item Type (centered) and their interaction, including by-participant and by-item random intercepts and a by-participant random slope for Item Type. There was a significant interaction between Item Type and Condition, b = -1.64, Wald 95% CI = [-2.61, -0.66], z = -3.30, p < .001. In the Partially Ambiguous condition, there was no significant difference between accuracy for disambiguated items (M = 52.9%, 95% CI = [46.0%, 59.9%]) and (partially) ambiguous items (M = 58.8%, 95% CI = [51.9%, 65.8%]), b = 0.43, Wald 95% CI = [-0.20, 1.07], z = 1.33, p = .18 (Figure 2.5C). However, in the Fully Ambiguous condition, participants more accurately produced the label for items disambiguated during training (M = 61.6%, 95% CI = [51.6%, 71.6%]) than for (fully) ambiguous items (M = 42.9%, 95% CI = [32.9%, 52.8%]), b = -1.20, Wald 95% CI = [-1.94, -0.47], z = -3.22, p = .001 (Figure 2.5B).

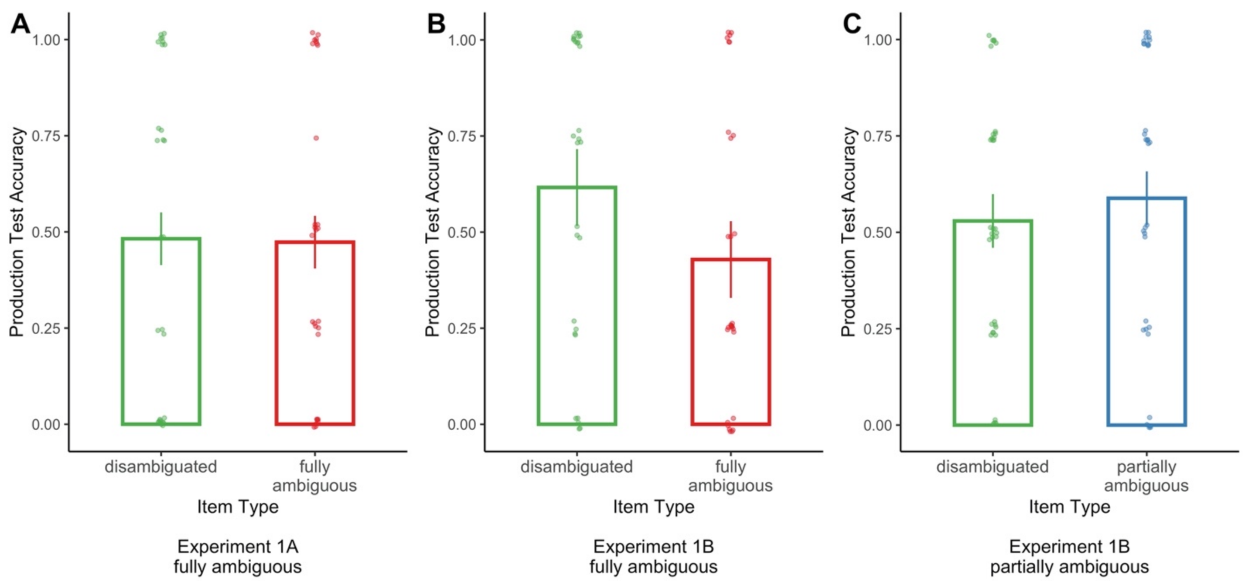

Figure 2.5. Production Test accuracy by item for (A) Experiment 1A, (B) the Fully Ambiguous condition in Experiment 1B, and (C) the Partially Ambiguous condition in Experiment 1B. Error bars represent within-participant 95% CIs (Morey, 2008).

Relationship between sampling and test performance

We next investigated the relationship between participants' sampling behavior and their subsequent performance at test. Specifically, we were interested in whether participants who preferentially sampled ambiguous items were more successful at learning the novel words.

Experiment 1A. Participants who chose more objects from the ambiguous set during the sampling phase accurately identified more words at test, $r(26) = .58$, 95% CI = [0.27, 0.78], $p = .001$ (Figure 2.6A).

Experiment 1B. In the Fully Ambiguous condition, participants' proportion of ambiguous selections was correlated with their test accuracy before the Sampling Phase (Test Block 1), $r(26) = .45$, 95% CI = [0.10, 0.71], $p = .015$. The proportion of ambiguous selections was marginally correlated with participants' test accuracy after the Sampling Phase (Test Block 2), $r(26) = .35$, 95% CI = [-0.03, 0.64], $p = .07$ (Figure 2.6B). In the Partially Ambiguous condition, participants' proportion of (partially) ambiguous items selected was not significantly correlated with their test accuracy before (Test Block 1: $r(32) = .09$, 95% CI = [-.24, .42], $p = .58$) or after the Sampling Phase (Test Block 2: $r(32) = -.12$, 95% CI = [-.44, .23], $p = .50$; Figure 2.6C).

By including a test phase immediately preceding the Sampling Phase, we aimed to further understand the correlation between test accuracy and preference for selecting ambiguous items observed in Experiment 1A. Specifically, do participants have higher test accuracy at the conclusion of the experiment because they preferentially selected ambiguous items, or do participants who are more successful at learning the object-label associations show a stronger preference for selecting ambiguous items?

To test this question, we correlated participants' proportion of ambiguous selections with their increase in accuracy from Test Block 1 to Test Block 2. If participants' ambiguous selections are driving higher accuracy, then participants who show a preference for sampling ambiguous items should show the largest increases in accuracy from Test Block 1 to Test Block 2. However, proportion of ambiguous items selected was not significantly correlated with an increase in test accuracy in the Fully Ambiguous condition ($r(26) = -.08$, 95% CI = [-.44, .30], p =.69) and negatively correlated in the Partially Ambiguous condition ($r(32) = -.36$, 95% CI = [-.62, -.03], p =.04), i.e. participants who were more likely to select the (partially) ambiguous items showed a lesser increase in test accuracy. Similar relationships were found with test accuracy for both ambiguous and disambiguated items.

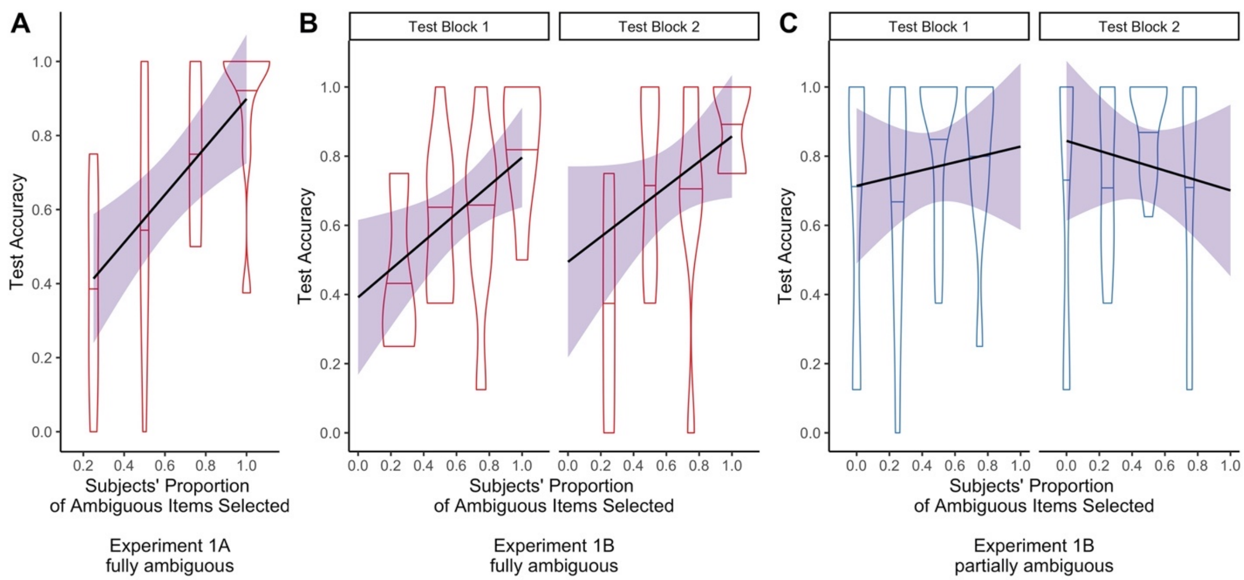

Figure 2.6. Relationship between choosing more ambiguous items and test accuracy for (A) Experiment 1A, (B) the Fully Ambiguous condition in Experiment 1B (split by Test Block), and (C) the Partially Ambiguous condition in Experiment 1B (split by Test Block). The distribution of individual participants' test accuracy for different proportions of ambiguous choices is represented with violin plots. Error bands represent +/-1 SE.

Discussion of Experiments 1A and 1B

In a cross-situational learning task, adult learners chose to learn more about those object-label pairs that remained ambiguous throughout training. Adults showed this tendency when the object-label pairings remained completely ambiguous based on the training evidence (Fully Ambiguous condition), but not when the object-label pairs became disambiguated at any point during training (Partially Ambiguous condition). Thus, this experiment provides evidence that adult learners will seek to reduce ambiguity about object-label associations when given the opportunity to control which items they will learn about.

At test, participants tended to show poorer overall learning of the ambiguous object-label pairs in the Fully Ambiguous condition, though not in the Partially Ambiguous condition. This finding is consistent with the fact that the fully ambiguous items were (by definition) designed to be more difficult object-label mappings to learn. However, it is important to note that a strong difference in accuracy between ambiguous and disambiguated items was found only in Experiment 1B, and not in Experiment 1A. The source of this discrepancy is not clear from the current results: besides simply being due to sampling error, these differences could be related to differences in experimental setting or slight differences in learning strategy. For instance, participants in the Fully Ambiguous condition in Experiment 1B not only showed slightly lower accuracy on ambiguous items, but also slightly better performance on disambiguated items compared to Experiment 1A. One possible explanation for this trend is that participants in Experiment 1B engaged in a more explicit strategy of disambiguating object-label pairings in the in-lab experimental setting, leading to more success at disentangling the object-label mappings that allowed for disambiguation. Past work on cross-situational learning has documented that explicit and implicit learning mechanisms interact during cross-situational word learning and are sensitive to shifts in task demands and task settings (Kachergis, Yu, & Shiffrin, 2014; Roembke & McMurray, 2016, 2020; Romberg & Yu, 2014, 2015). However, regardless of whether participants engaged in slightly different learning strategies in Experiment 1A and 1B, they consistently showed a preference for sampling ambiguous items in the Fully Ambiguous condition.

Intriguingly, we found that participants' sampling behavior was correlated with their test accuracy in the Fully Ambiguous condition: participants who chose more ambiguous items during the Sampling Phase also more accurately identified object-label associations. By testing participants before and after the Sampling Phase in Experiment 1B, we were able to partially disentangle the directionality of this effect. Participants who had learned the novel words better following the Training Phase were more likely to sample ambiguous items. In other words, participants' learning success appeared to predict their likelihood of targeting ambiguous items during the Sampling Phase, rather than participants' sampling choices (solely) driving their test accuracy. Perhaps surprisingly, participants' likelihood of selecting ambiguous items did not appear to lead to bigger increases in accuracy from the first to the second testing (though accuracy for ambiguous items showed the largest increase in general). There are a few potential reasons why we did not observe greater increases in accuracy following ambiguous item selections in the present study. First, participants who showed a strong preference for selecting ambiguous items were already performing quite well at test, suggesting that the lack of an effect may in part be due to a ceiling effect. Second, the Sampling Phase was designed such that participants continued to receive ambiguous, cross-situational learning trials: even after selecting ambiguous items, participants would view a cross-situational learning trial involving an additional, randomly selected item. While these trials in principle provided the opportunity to disambiguate object-label mappings, they also required participants to simultaneously continue to update and maintain the object-label relationships from past training, which may in turn have "washed out" some of the learning benefits that participants might have accrued from targeting the ambiguous items during the Sampling Phase. Future research could specifically target what types of sampling behavior leads to better learning outcomes, relative to randomly selected training events.

Experiment 2A

Experiments 1A and 1B established that adult learners seek to reduce uncertainty about ambiguous items when learning novel words. Next, we asked whether children would demonstrate a similar tendency to seek new words that reduce ambiguity during cross-situational learning. As in Experiment 1A, children (4 - 8 years of age) completed a cross-situational word learning task. Across training, one set of novel object-label associations could be inferred based on the object-label associations they co-occurred with, while another set of words remained ambiguous. Then, participants were given the opportunity to sample object-label associations presented in isolation, i.e. in unambiguous learning trials. The central question was whether children would prefer to select object-label associations with ambiguous evidence during training, suggesting that children sample words that reduce referential ambiguity.

Method

Participants

We recruited 38 participants (M = 5.9 years, SD = 1.19, range: 4.1 – 8.1 years, 19 female) at a local children's museum. Two additional participants were excluded due to inattention during the experiment.

Stimuli

The object stimuli were 8 images of novel 'alien' creatures used in previous word learning studies (Partridge et al., 2015) and 2 cartoon images of familiar animals (penguin, dog). 8 novel word stimuli (*biffer, deela, guffi, sibu, tibble, leemu, zeevo, pahvy*) and two familiar word stimuli (*penguin*, *dog*) were recorded by a female native speaker of English and normalized in duration and average loudness. The association between each novel label and its novel target referent, as well as the particular roles of the novel word-referent stimuli, were randomized across participants. The stimuli were presented using in a web-based experiment created in jsPsych (de Leeuw, 2015).

Design & Procedure

Children were tested in a quiet room in the children's museum on a 10.1" Samsung Galaxy Note tablet. An experimenter guided children through the task by giving instructions at the beginning of each new phase. The experiment was presented as a game in which a cartoon bear named Teddy would first teach children the names of new alien friends, and then ask children to help her find her friends. The experimenter began with the following introduction:

> In this game, Teddy went up to space and met a bunch of new alien friends. Teddy is going to tell you the names of aliens, and your job is to try to remember which name goes with which alien. Later, you're going to help Teddy find them.

The experiment then proceeded to a Practice Phase, followed by the main experiment consisting of three phases: Training, Sampling, and Test.

Practice Phase. Participants first completed a practice phase in which they encountered the two familiar word object stimuli and two novel object-label associations. We introduced this short practice phase to give children experience with the overall structure of the main experiment under less demanding circumstances, using a smaller set of items and mixing familiar and novel items. First, children were exposed to 4 practice training trials similar in structure to the training trials in the main experiment. On each trial, two referents appeared on the screen on either side of the Teddy character and children heard two labels, one for each object, in random order. On the first trial, children always saw the two familiar items (i.e., the penguin and the dog), followed by a second trial in which children saw two novel object-label associations (i.e., an ambiguous labeling event). On the final two training practice trials, children saw each of the familiar items occur with one of the two novel items (permitting the disambiguation of the novel object-label associations). Next, children saw two sampling practice trials, in which they had the opportunity to select which of the four items they wanted to learn about next, followed by four practice test trials, in which participants' knowledge of the items was tested in a 4-AFC recognition test. The procedure for each of these practice trial types mirrored the procedure for the Sampling phase and the Test phase described in more detail below.

Training Phase. Participants completed 9 cross-situational learning trials (3 blocks of 3 trials each). On each training trial, participants saw two referents appear on the screen on either side of the Teddy character and heard the labels of the two objects presented sequentially in random order. Next, the objects switched locations in a brief animation, and participants heard the same two labels presented in the same order. We introduced this trial repetition with flipped locations in order to reduce children's tendency to interpret the labeling event as moving from left to right on the screen, i.e. assuming that the first label went with the object on the left and the second label went with the object on the right.

As in the Fully Ambiguous condition of Experiment 1A, we manipulated whether the object-label associations could be disambiguated across trials during training (Figure 2.7). Every object-label pair occurred on three cross-situational training trials. Four of the objects occurred with three different object label pairs (disambiguated items). The remaining two object-label associations always occurred with one another (ambiguous items), such that children never saw evidence allowing them to link the two words unambiguously with their respective referent.

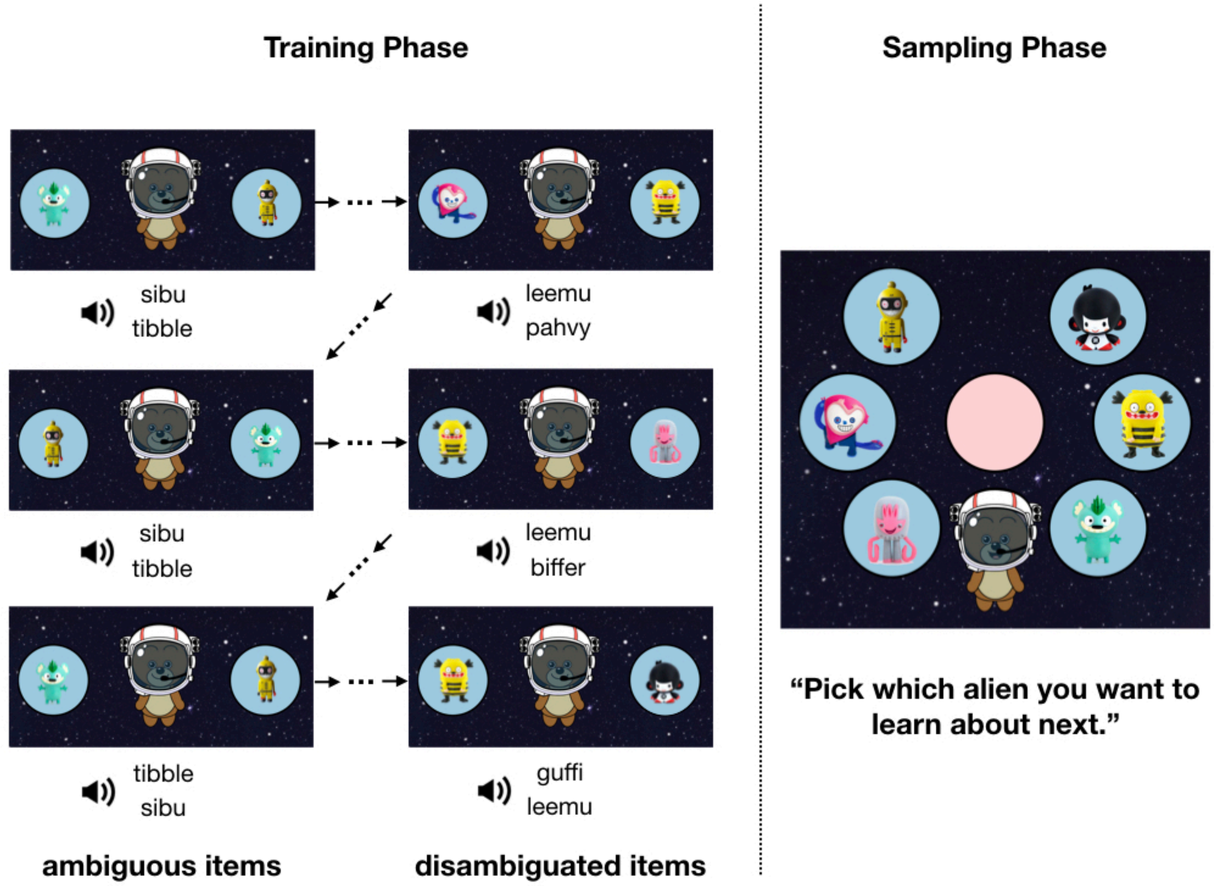

Figure 2.7. Overview over the design of the Training and Sampling Phase in Experiment 2A.

Sampling Phase. After completing the training phase, participants completed four sampling trials. On each sampling trial, all 6 referents appeared in randomized locations on the screen. Participants were instructed to select which of the 6 items they wanted to learn about next (Figure 2.7). When participants tapped one of the 6 referents, a brief animation moved the item to the center of the screen while the remaining items disappeared, and the referent was subsequently labeled in isolation.

Test Phase. Participants' knowledge of the object-label associations was probed in a 6-AFC recognition test. On each test trial, all 6 referents appeared in randomized locations on the screen surrounding the Teddy character. When participants tapped Teddy in the center of the screen, they heard one of the 6 labels. Participants were instructed to help Teddy by selecting the friend she was looking for. No feedback was provided after a choice. Participants were tested on each label in random order, for a total of 6 recognition test trials.

Predictions

As in Experiment 1A, our main prediction was that children would preferentially select object-label associations that remained ambiguous during the cross-situational word learning trials of the training phase.

Results

Sampling choices

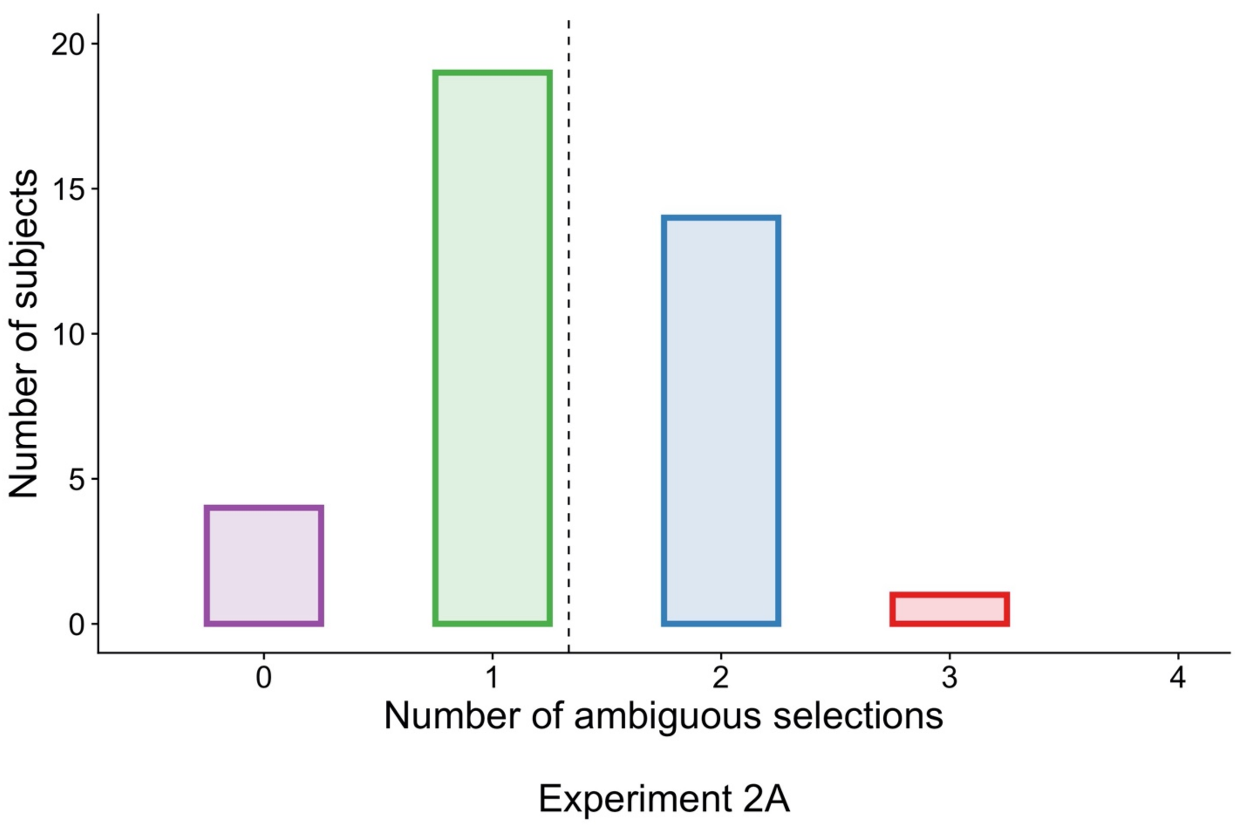

Figure 2.8. Children's sampling choices in Experiment 2A. The plot depicts the number of subjects (out of 38) selecting 0, 1, 2, 3, or 4 ambiguous items across the four sampling trials. The dashed line represents the expected average value if items are sampled randomly.

We fit a logistic-mixed effects model testing whether children's likelihood of selecting an ambiguous item differed from chance (logit(0.33)), including by-participant and by-item random intercepts. Contrary to our prediction, children did not preferentially select ambiguous object-label associations during the Sampling phase, b = -0.02, 95% CI = [-0.36, 0.32], z = -.12, p = .91. Participants chose an object from the ambiguous set on 32.9% of trials (95% CI = [27.1%, 38.7%]; see Figure 2.8.). A Wilcoxon signed-rank test conducted on children's proportion of ambiguous selections yielded comparable results (V = 409, p = .57).

Test performance

Overall, participants showed significant learning of the object-label pairs, choosing the correct object to go with a label at above-chance levels (chance = 0.167), M = 38.6%, 95% CI = [30.7%, 46.5%], $t(37) = 5.65$, $p < .001$. However, surprisingly, children performed more accurately on the ambiguous items (M = 48.6%, 95% CI = [36.9%, 60.4%]) than on the disambiguated items (M = 33.6%, 95% CI = [24.9%, 42.2%]; Figure 2.9). This difference was significant in a logistic mixed-effects model predicting trial-by-trial accuracy from Item Type (centered; ambiguous = 0.5; disambiguated = -0.5), including by-participant and by-item random intercepts, and a by-participant random slope for Item Type, b =.68, 95% Wald CI = [0.08, 1.28], $z = 2.23$, $p = .026$. When tested on ambiguous items, children had a strong preference to select one of the two ambiguous objects (61.8% of trials, 95% CI = [50.7%, 72.9%]) rather than one of the four disambiguated objects (chance = 0.33). When tested on disambiguated items, children tended not to choose the two ambiguous objects, selecting them on only 18.4% of trials (95% CI = [12.8%, 24.1%]).

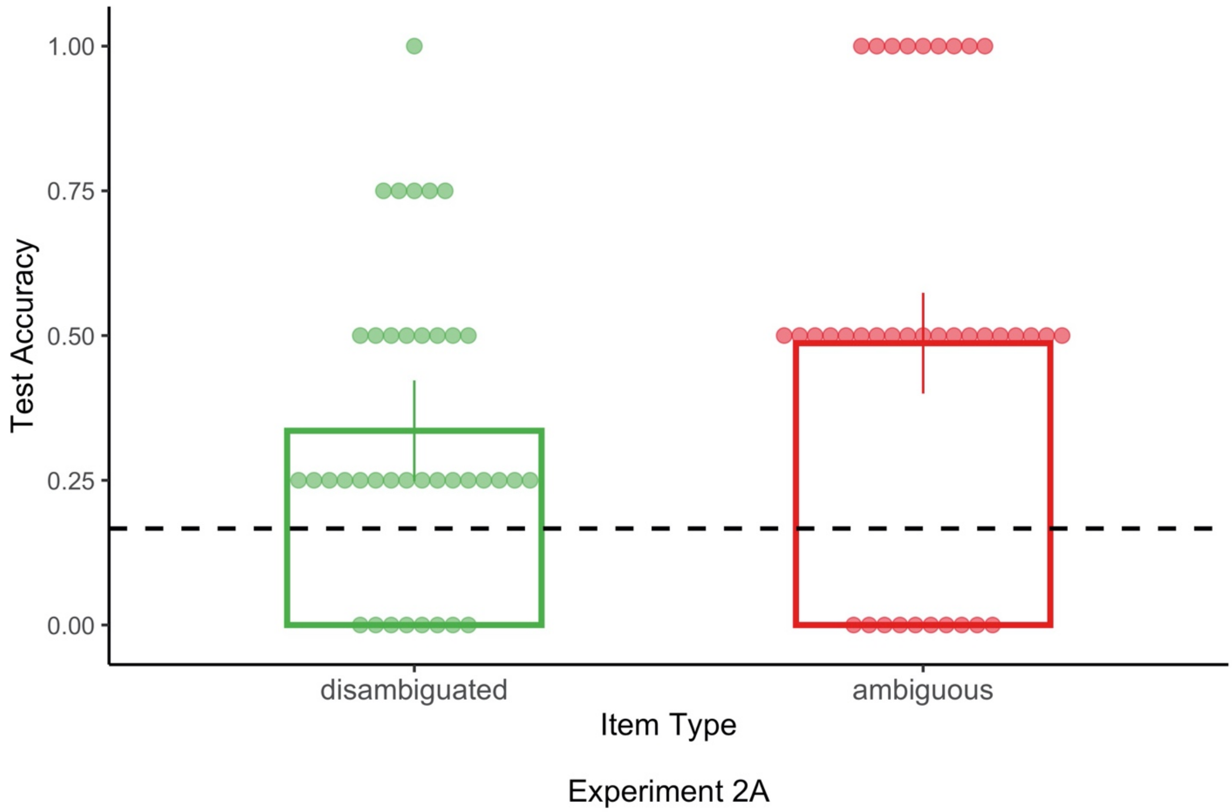

Figure 2.9. Children's test accuracy in Experiment 2A, split by Item Type (disambiguated vs. ambiguous). Error bars represent within-participant 95% CIs (Morey, 2008).

Relationship between sampling selections and test performance

We next investigated the relationship between children's selections during the Sampling Phase and their subsequent accuracy on sampled (vs. non-sampled) items. To test the impact of children's selections on learning, we fit a logistic mixed-effects model predicting children's test accuracy from Item Type (centered; ambiguous = 0.5; disambiguated = -0.5), Sampling Choice, i.e. whether or not the item was chosen by a participant during the Sampling Phase (centered; sampled = 0.5; not sampled = -0.5), and their interaction, including by-participant and by-item random intercepts, and a by-participant random slope for Item Type. There was a significant effect of Item Type (as above), b =.83, 95% Wald CI = [0.15, 1.50], z = 2.40, p = .016. There was also a significant effect of Sampling Choice, b =.89, 95% Wald CI = [0.23, 1.55], z = 2.66, p

= .008 (Figure 2.10), indicating that participants performed more accurately at test on items that they selected during the Sampling Phase (M = 45.2%, within-participant 95% CI = [34.6%, 55.7%]) than on items they did not sample (M = 26.8%, within-participant 95% CI = [16.2%, 37.3%]). There was no significant interaction between Item Type and Sampling Choice, p = .48.

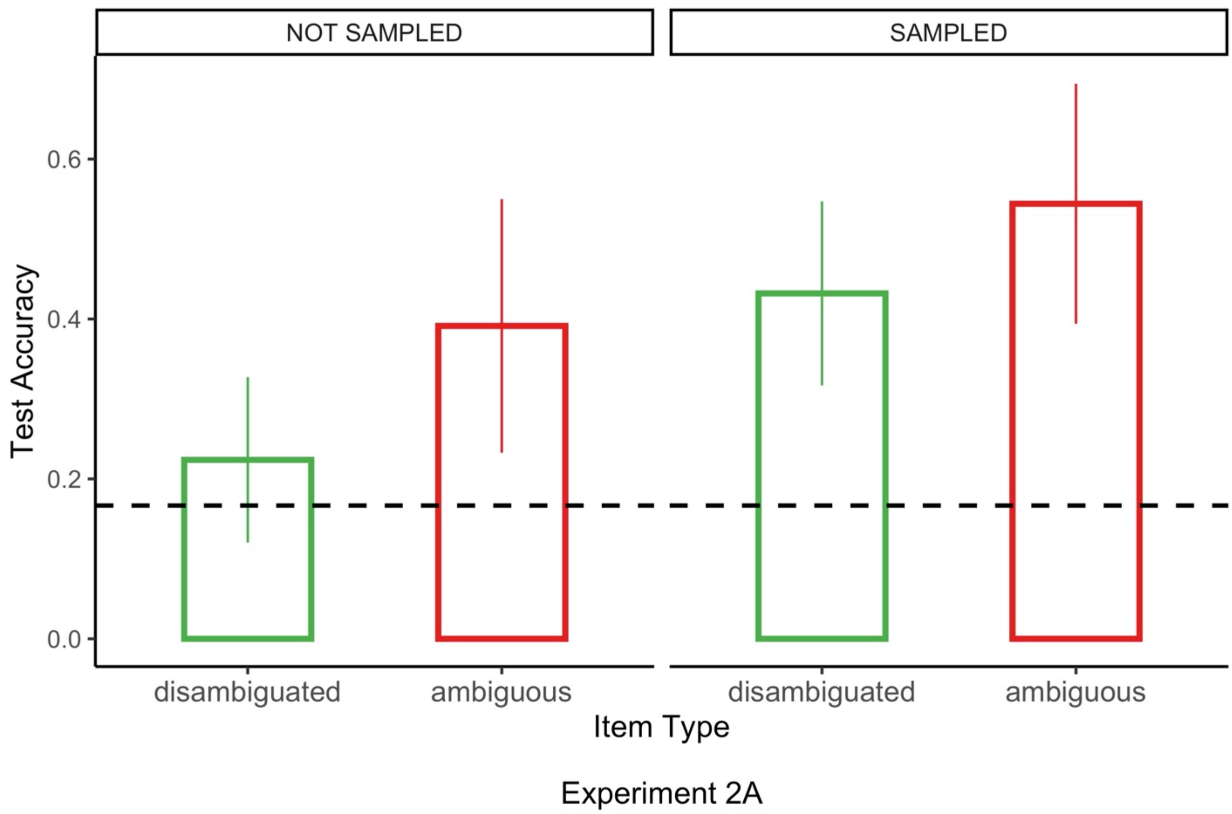

Figure 2.10. Children's test accuracy in Experiment 2A, split by Item Type (disambiguated vs. ambiguous) and whether a given test item was selected during the Sampling Phase (not sampled vs. sampled). Error bars represent within-participant 95% CIs (Morey, 2008).

Discussion

Unlike adult learners, children did not show a preference for selecting object-label associations for which they had experienced ambiguous evidence during training. Interestingly, children performed better on items testing ambiguous object-label associations than for object-label associations that were disambiguated across training trials. There are likely two reasons why children showed higher accuracy on the ambiguous items. First, since the two ambiguous items always co-occurred with one another, the training could help learners constrain the set of possible competitors for a given ambiguous label to two objects (compared to four possible objects for the disambiguated items). Indeed, children appeared to constrain their choices to the two objects that co-occurred on ambiguous trials when tested on their respective labels and rarely chose these objects when tested on the labels that occurred with the disambiguated objects.

Second, anecdotally, we observed that many children explicitly pointed to specific objects during training while listening to each label and even repeated the respective label for each object. This behavior may indicate that some children were making an explicit hypothesis about each word mapping (Trueswell, Medina, Hafri, & Gleitman, 2013). If a child formed a specific hypothesis about the mapping between the two labels and objects on the first ambiguous trial, they would subsequently hear evidence that would appear to confirm their hypothesis: the two labels and the two objects would occur together again on the subsequent two training trials. "Hypothesis-testers" would never experience evidence disconfirming their initial hypotheses and thus have a 50% chance of responding correctly at test for these items (note that our participants' test accuracy was 48.6% on average). Crucially, one consequence of learners approaching the task in this manner is that the two object-label associations deemed "ambiguous" according to the experimental design may have actually appeared *less* ambiguous to children performing the task than the putatively disambiguated items. Thus, in our next step, we adapted the task to create a learning situation in which one set of object-label associations would be more clearly ambiguous from the standpoint of the child learner.

Experiment 2B

In Experiment 2B, we sought to increase the likelihood that children would perceive some novel object-label associations as more ambiguous than others. We used mutual exclusivity to increase the ease with which children could infer word-referent pairs for one set of novel objects (Markman & Wachtel, 1988) while maintaining the ambiguity of a second set of novel word-referent pairs as in the previous experiments. By giving children the opportunity to infer the referents for novel objects occurring in mutual exclusivity trials, we aimed to make it easier for children to recognize the referential ambiguity of novel object-label associations that always co-occurred.

Method

Participants

We recruited 56 participants (M = 5.5 years, SD = 1.18, range: 3.3 – 7.9 years, 33 female)[4] at a local children's museum. Two additional participants were excluded due to interruptions to the experiment (n = 1) or for not completing the study (n = 1).

Stimuli

The novel object and word stimuli were six images and recordings composed of a subset of the items used in Experiment 2A. In addition, 4 cartoon images of familiar animals (cow, dog, monkey, pig) along with audio recordings of their respective labels were used. All word stimuli were recorded by the same female native speaker of English and normalized in duration and average loudness. A demo of the experiment can be accessed through a web browser at

---

[4] The original target age range for the study was 4.0 years -8.0 years. Three children below the age of four years (all 3-year-olds) were recruited and run in the experiment. Given that all three children completed the experiment without issue, we opted for an inclusive data policy and included these participants in the current analysis. All analytic results and conclusions are similar if these three participants are excluded.

http://sapir.psych.wisc.edu/~zettersten/demos/Crossact_kids/crossact_v4.html (note that trials are typically advanced by clicking on the Teddy character appearing in the center of the screen).

Design & Procedure.

The procedure and testing conditions were identical to Experiment 2A. The experiment followed the same structure as Experiment 2A, beginning with a *Practice Phase* and then proceeding through the *Training Phase*, *Sampling Phase*, and *Test Phase*.

Training Phase. Participants completed 9 cross-situational learning trials (3 blocks of 3 trials each) with 6 object-label pairs, two familiar object-label pairs (e.g., pig and dog) and four novel object-label pairs chosen randomly from the set of novel stimuli. As in Experiment 2A, on each trial, participants saw two referents appear on the screen and heard two labels presented in random order. Two novel object-label associations always occurred with one another (ambiguous items), mirroring the ambiguity manipulation from Experiments 1A/B and 2A. The two remaining novel object-label associations were each yoked to one of the two familiar object-label pairs (i.e., one alien always occurred with the dog image, while the other always occurred with the pig image; mutual exclusivity items). We reasoned that children would successfully disambiguate reference for mutual exclusivity items (i.e., on seeing an image of a dog and a novel "alien" and hearing the words *leemu* and *dog*, children would successfully infer that *leemu* referred to the novel alien). This would make it more likely that the ambiguous items would be perceived by child learners as having high referential uncertainty. As in previous experiments, all novel objects and their labels occurred equally frequently across the training phase.

Sampling Phase. Participants next completed two sampling trials. On each trial, the four novel objects appeared on the screen and children were instructed to choose which object they wanted to learn more about. The procedure was otherwise identical to Experiment 2A.

Test Phase. Participants' knowledge of the six words from the training phase (4 novel, 2 familiar words) was tested in a 6-AFC recognition task as in Experiment 2A.

Results

Sampling choices

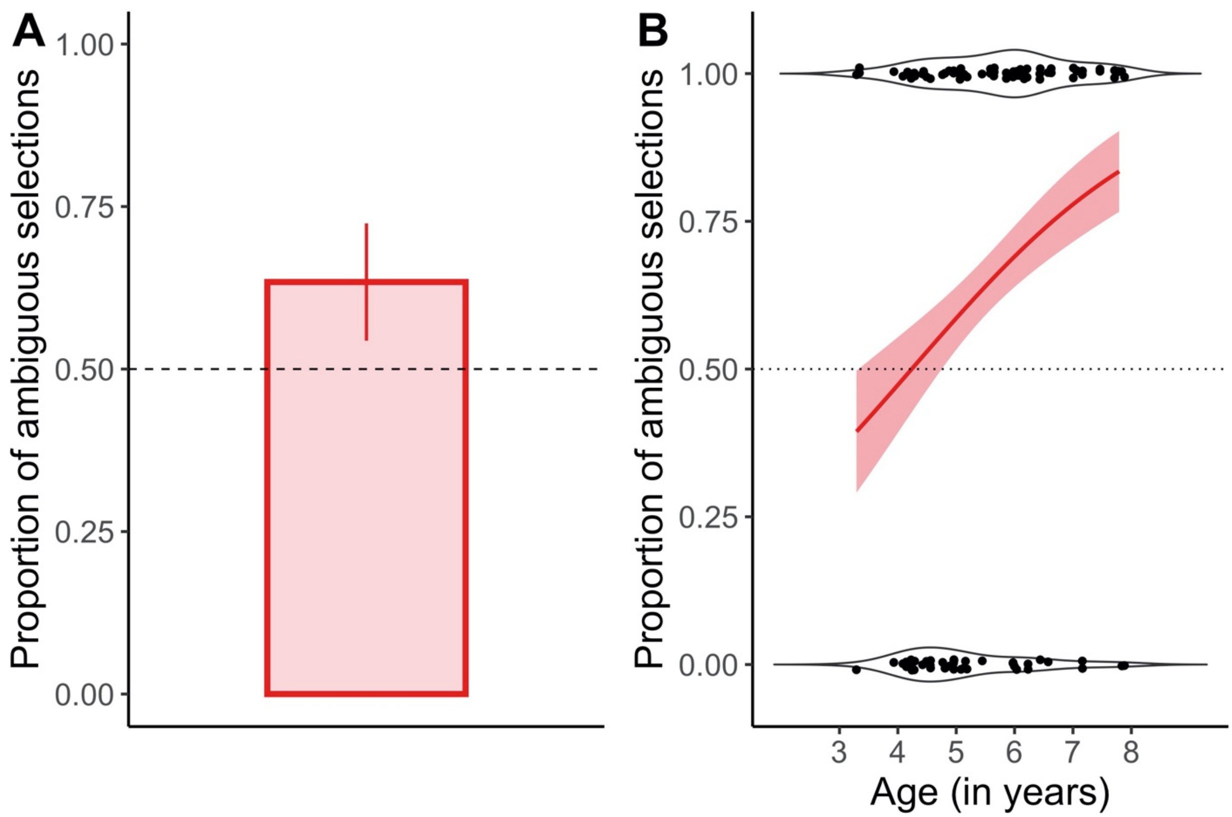

Figure 2.11. Proportion of ambiguous item selections in Experiment 2B overall (A) and across age (B). Error bars represent 95% CIs and error bands are +1/-1 SEs based on model estimates.

Children preferentially selected ambiguous object-label associations during the Sampling Phase, $b = 0.55$, Wald 95% CI = [0.15, 0.95], $z = 2.71$, $p = .007$. Participants chose an object from the ambiguous set on 63.4% of trials (95% CI = [54.4%, 72.4%]) (chance level = 0.5; Figure 2.11A). A Wilcoxon signed-rank test yielded similar results ($V = 330$, $p = .006$). The likelihood of children making ambiguous selections increased with age, $b = 0.45$, Wald 95% CI

= [0.10, 0.82], $z = 2.48$, $p = .013$ (logistic mixed-effects models including Age as a fixed effect; Figure 2.11B).

Test performance

Overall, participants showed significant learning of the object-label pairs, choosing the correct object to go with a label at above-chance levels (chance selection of novel object = 0.25), M = 57.6%, 95% CI = [48.4%, 66.8%], b =1.53, z = 5.08, p < .001. Accuracy for mutual exclusivity items (M = 64.2%, 95% CI = [53.2%, 75.1%]) and for the ambiguous items (M = 55.7%, 95% CI = [44.0%, 67.3%]) was similar, b = -.44, Wald 95% CI = [-1.15, 0.27], z = -1.22, p = .22 (Figure 2.12).

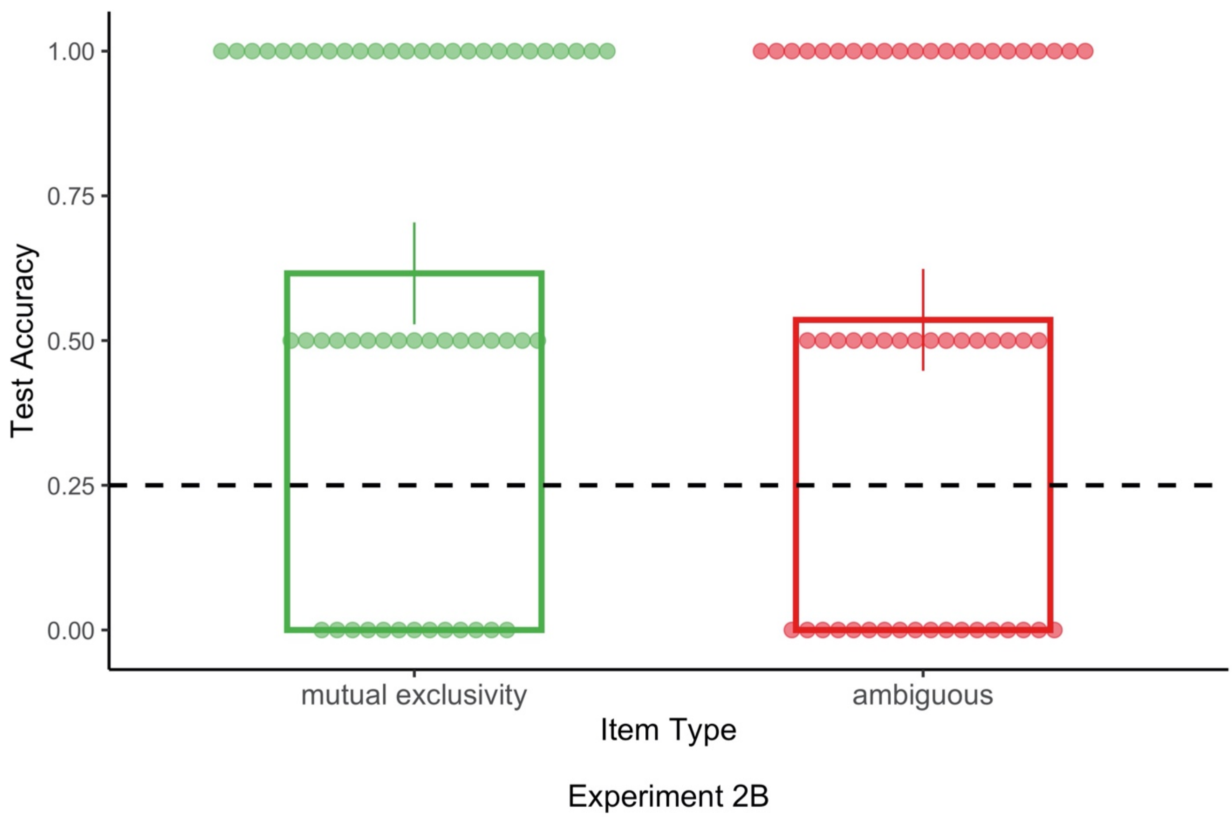

Figure 2.12. Children's test accuracy in Experiment 2B, split by Item Type (disambiguated/ mutual exclusivity vs. ambiguous). The dashed line represents chance level for selecting novel objects. Error bars represent within-participant 95% CIs (Morey, 2008).

Relationship between sampling selections and test performance

As in Experiment 2A, we investigated the relationship between children's selections during the Sampling Phase and their subsequent accuracy on sampled (vs. non-sampled) items. We fit the same logistic mixed-effects model predicting children's test accuracy from Item Type (centered; ambiguous = 0.5; mutual exclusivity = -0.5), Sampling Choice, i.e. whether or not the item was chosen by a participant during the Sampling Phase (centered; sampled = 0.5; not sampled = -0.5), and their interaction, including by-participant and by-item random intercepts, and a by-participant random slope for Item Type. There were no significant effects of Sampling Choice (p = .33) or Item Type (p = .15), and no significant interaction between the two (p = .77).

Discussion

When given the opportunity to select which object-label pairs they wanted to learn more about, 3-8-year-olds preferentially selected object-label pairs that remained ambiguous during training over object-label pairs that could be disambiguated through mutual exclusivity. These findings demonstrate that – at least in some ambiguous word learning situations – children prefer to select learning events that aid in reducing referential uncertainty. The tendency to make ambiguity-reducing selections began to emerge around 5 years of age in our sample.

General Discussion

When learning the referents of novel labels in ambiguous contexts, adult learners chose to learn more about object-label associations that remained more ambiguous at the end of training. These choices were related to participants' learning: participants who were most successful at learning object-label associations during training were also most likely to systematically select more ambiguous items during the Sampling Phase. It is interesting to note the modest magnitude of adults' preference on the task: ambiguous items were selected on slightly less than two-thirds of adults' sampling trials. This may be partly related to the design of the sampling phase, which allowed for a number of potentially successful sampling strategies (e.g., selecting a known word on each sampling trial in order to hear that known word in combination with other words). However, another intriguing possibility for future research is that there are individual differences in how adults organize their learning, and that these differences may lead to distinct learning outcomes (see also Kachergis et al., 2013).

Children also spontaneously sampled object-label associations that reduced ambiguity, though only when the task was simplified to emphasize referential ambiguity. When presented with a similar task as adults, 4-8-year-olds did not choose to learn about object-label associations that remained ambiguous during training. However, this result is likely at least partially explained by the fact that children actually found disambiguated items *more* difficult to learn than ambiguous items. In a simplified design that highlighted the ambiguous nature of the trials in which two referents always occurred together, children chose to learn about items that reduced uncertainty about the words' referents.

The preference for selecting ambiguous items was strongly related to age, with children beginning to reliably select the ambiguous items around 5 years of age in our sample. Past work on social referencing suggests that children as young as 2 years of age (Hembacher & Frank, 2017) and even infants as young as 12 months are sensitive to referential uncertainty (Bazhydai et al., 2020; Vaish et al., 2011). Our studies go beyond measuring sensitivity to uncertainty by asking whether child learners will explicitly make decisions to sample new words based on referential ambiguity. Proactively making sampling decisions based on uncertainty may require more sophisticated skills in metacognition (Ghetti, Hembacher, & Coughlin, 2013; Lyons &

Ghetti, 2011) and cognitive control (Munakata, Snyder, & Chatham, 2012) that undergo substantial development during early childhood. The limits on the extent to which younger children spontaneously make ambiguity-reducing selections raise important questions for future research about how children's sampling strategies develop and interact with the cognitive development more generally.

Children have substantial control over their "curriculum" as they learn new words in the world (Mani & Ackermann, 2018; Smith et al., 2018), with potentially immense consequences for the difficulty of the learning problem they face (Hidaka et al., 2017). The present results demonstrate that, at least in some circumstances, children will sample new words that reduce referential ambiguity. These studies contribute to a growing literature demonstrating that children are curious learners who actively contribute to their own language development.

Chapter 3: Does active sampling support learning new words?

The studies in chapter 2 demonstrated that, at least under some circumstances, both children and adults sample information that is supportive of learning object-label associations, in the sense of reducing uncertainty or ambiguity about new word meanings. An outstanding question from these findings is what type of consequences children's sampling has for learning. Do children learn new words better if they can actively control their learning input than if they cannot?

Past research has established that having active control over the learning input supports some forms of category and word learning (Castro et al., 2009; Kachergis et al., 2013; Markant & Gureckis, 2014). In a groundbreaking study, Markant & Gureckis (2014) found that adult participants were better able to learn simple category rules when given the ability to control which category exemplars they could learn about (active/ self-directed condition), compared to two passive learning conditions: a passive condition in which participants saw a random sample of category exemplars and a passive yoked condition in which participants could not control their learning input, and instead viewed the same category exemplars generated by a participant from the active condition. In the domain of word learning, Kachergis et al. (2013) found that allowing participants to construct their own learning input during an ambiguous word learning task, by selecting which sets of items would occur in a cross-situational word learning task, helped participants learn novel names better than participants who saw randomly generated trial lists. At least in certain contexts, controlling one's own input helps adult learners acquire novel categories and words more effectively.

Why does active control support learning? Two main explanations have been advanced in past literature: hypothesis-dependent sampling and memory-enhancing mechanisms (Gureckis &

Markant, 2012; Markant, Ruggeri, Gureckis, & Xu, 2016). One reason why self-directed sampling boosts learning is because it allows learners to select material that helps adjudicate between hypotheses they are currently entertaining (Markant & Gureckis, 2014; Markant, Settles, & Gureckis, 2016). For example, in a word learning task such as the one used by Kachergis et al. (2013), learners in the active condition can specifically select those objects for which they are currently considering different possible names to occur on their next training trial. On the other hand, participants in a passive condition cannot match their next training sample to the hypotheses about word meanings they are currently considering, making it more difficult to update their hypotheses. A second explanation appeals to the diverse means in which active control can boost learners' memory (Markant, Ruggeri, et al., 2016). For example, being able to actively control the learning process may lead to higher engagement and arousal during training (Berlyne, 1960; Kidd & Hayden, 2015), the added cognitive demands of planning goal-directed selections could lead to greater elaboration of the learning materials (Kachergis, Rhodes, & Gureckis, 2017), and active control could aid learners in coordinating selective attention towards the to-be-encoded material (Markant, DuBrow, Davachi, & Gureckis, 2014).

While there is ample support for the finding that active control supports category and word learning in adults, the evidence on active learning benefits in children is less extensive and more mixed. While there is some evidence from tasks inspired by the category learning work of Markant and Gureckis (2014) that children can induce a novel category boundary better when allowed to actively sample category exemplars (Adams et al., 2017; Sim, Tanner, & Alpert, 2015), effects appear to often be small, with high individual variability. Little is known about whether active control supports word learning in children. One previous word learning study (N = 32) has reported active learning benefits for children around the age of 3 (Partridge et al., 2015). However, the benefit of active learning was only found for small sets of objects (1-2 novel words) and was only compared to a passive condition in which items were selected at random. Moreover, no study to date has investigated children's word sampling choices in relation to learning outcomes, and whether selection strategies explain differences in outcomes between active and passive learning conditions.

Experiment

We asked whether allowing children to actively control a portion of their training input during an explicit word learning task would support novel word learning. To test this question, we compared children's word learning in an Active condition, in which they could choose novel words to learn about next, to their performance in two yoked passive conditions: the Yoked Passive condition and the Yoked Passive Exposure Mismatch condition. All children were first given initial exposure to a novel set of object-label associations. Children's subsequent training exposure differed depending on the sampling condition. In the Active condition, children were given the opportunity to select which novel-object label associations they would hear next. In the Yoked Passive condition, children were passively exposed to – but did not select - the same novel labels seen by a (yoked) participant from the Active condition. The Yoked Passive Exposure Mismatch condition was identical to the Yoked Passive condition, except that the yoked participant from the Active condition had experienced an initial labeling exposure that differed from the child in the passive condition. In other words, children in the Yoked Passive Exposure Mismatch condition and their yoked active counterpart saw the same input during the sampling phase of the experiment, but were considering this input under different initial learning experiences. This condition originally arose as an error in the programming of the task – however, the resulting data allowed us to investigate to what extent benefits from active learning can arise from children's ability in the Active condition to tune their sample to their past learning experience.

Method

Participants

We recruited 231 participants (M = 4.6 years, SD = .84, range = 3.0 – 6.0 years, 121 female) at a local children's museum. 40 additional participants were excluded due to falling outside the age range of the study (n=4), experimenter error (n=5), equipment failure (n=1), due to a developmental concern (n=2), for not completing the study (n=16), for having participated in an earlier version of the study (n=10) or due to parental or sibling interference (n=2). Participants were assigned to one of three conditions: the Active condition (n=77), the Yoked Passive condition (n=77), or the Yoked Passive Exposure Mismatch condition (n=77). Since participants in the two Yoked Passive conditions were yoked to participants from the Active condition, true random assignment was not possible. Assignment to the three conditions occurred as follows. The Active condition and the Yoked Passive Exposure Mismatch condition were collected together. For each age group (3-year-olds, 4-year-olds, and 5-year-olds), we first assigned a given participant to the Active condition. The next child belonging to the same age group was assigned to the Yoked Passive Exposure Mismatch condition, and yoked to the preceding child from the same age group assigned to the Active condition. Within each age group, we then continued to alternate between the Active condition and the Yoked Passive Exposure Mismatch condition, always yoking the next participant in the passive condition to the preceding child in the Active condition belonging to the same age group. After completing data collection, we determined that a programming error had led to a mismatch between the exposure for children in the Active condition and their yoked counterparts in the Yoked Passive Exposure Mismatch condition (see below for further details on why data from this condition was retained). We therefore collected an additional sample of children assigned to the Yoked Passive condition (in which this error was corrected). Since this condition was collected after the Active condition was completed, children were randomly assigned to a yoked counterpart from the Active condition belonging to the same age group.

Stimuli

The object stimuli were eight images of novel 'alien' creatures drawn from the same set of items used in previous word learning studies (Partridge et al., 2015) and four cartoon images of familiar animals (a bear, a cow, a penguin and a pig) (see Appendix A, Figure A4). Eight novel word stimuli (*beppo, finna, kita, manu, noopy, roozer, soma, tesser*) and four familiar animal words (*bear*, *cow*, *penguin*, *pig*) were recorded by a female native speaker of English and normalized in duration and average loudness. The association between the eight novel labels and the eight referent images was randomized across participants. The stimuli were presented using a web-based experiment created using jsPsych (de Leeuw, 2015). A demo of the experiment can be accessed through a web browser at http://sapir.psych.wisc.edu/~zettersten/demos/act/active.html.

Design & Procedure.

Children were tested in a quiet room in the children's museum on a 10.1" Samsung Galaxy Note tablet. A trained experimenter introduced children to the game and guided children through the task by giving verbal instructions. All participants were introduced to the experiment as a game in which they had to learn the names of the friends of a cartoon bear character. The experiment consisted of three blocks: a practice block and two main experimental blocks. Each block was composed of three phases: an Exposure Phase, a Sampling Phase, and a Test Phase. In the practice block, children were familiarized to the word learning game using the four known animal words and images. The goal of the practice block was to ensure that children understood the structure and the goal of the game. Children then completed the two main experimental blocks. In each block, children learned and were tested on four novel words. Each of the two main blocks were identical in structure (i.e., children completed the same task), but with two non-overlapping sets of four novel object and word stimuli. We included two identical experimental blocks to investigate whether there were learning effects in the second block in children's sampling behavior after children's experience in the first experimental block (e.g., to investigate whether children engaged in more target sampling behavior after having played the game once and thus having more experience with the task).

Exposure Phase. In the Exposure Phase, children were first given initial experience with the object-label associations. At the beginning of the Exposure Phase, four opaque circles appeared on either side of the Teddy character, along with a white square positioned in the center between the four circles. Next, each circle would turn red one at a time in random order, and children were instructed to tap each red circle to reveal one of the four images (either a familiar animal in the practice block or a novel alien creature in the experimental blocks). This ensured that children attended to each of the four images and had practice tapping each of the four image locations. After the four images were revealed, the center square turned red, and children were instructed that Teddy would tell them the names of her friends. On each trial, children tapped the red square to initiate a labeling event (see Figure 3.1). After touching the red square, one of the four images wiggled (in order to draw children's attention to the image) and moved to the center square, while all of the remaining objects were obscured by the four circles turning blue. Once the image was in the central square and all other objects were hidden (a procedure that took 1600 ms), an audio stimulus labeled the object using one of two randomly selected carrier phrases (*It's* *a* ____ or *That's a* ____). The auditory stimulus played for 1450 ms, after which the trial paused for 350 ms before ending. The next trial began with all four objects visible in the same location as before (i.e., the object locations stayed consistent across Exposure trials) and a red square that children were instructed to tap to initiate the next labeling event. The labeling events were designed to ensure that children's attention was focused on the labeled image by (a) drawing children's attention to the object in an initial animation, (b) presenting the to-be-labeled object in the central square, and (c) removing the other objects from view.

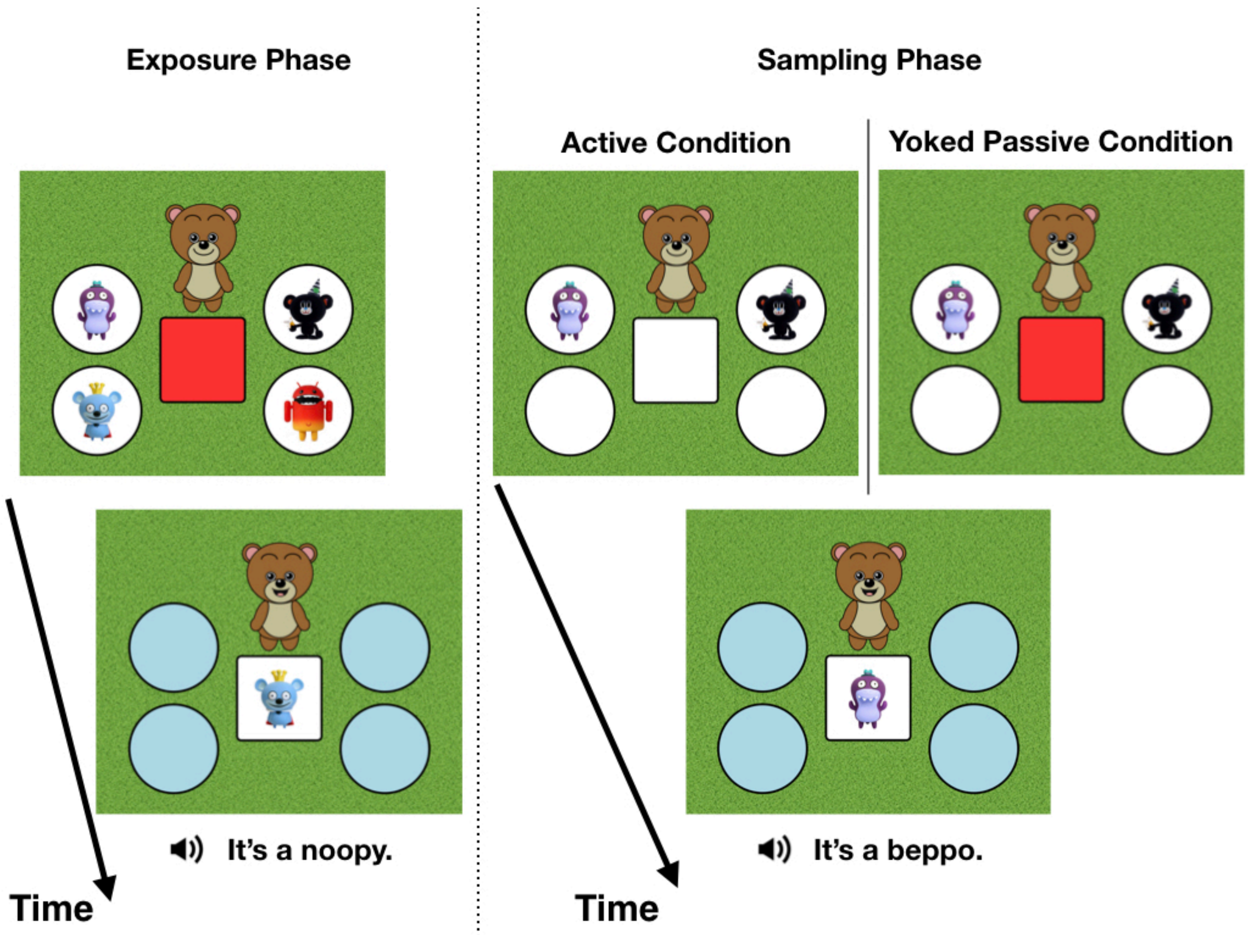

Figure 3.1. Overview over the design of the Exposure and Sampling Phase (Active and Yoked Passive Condition).

In the practice block, the four images were the four familiar animal images. Participants were presented with two exposure trials (consistently *bear* and *cow*). In the experimental blocks, children were presented with novel object-label associations in eight exposure trials. The key manipulation was the frequency with which children were exposed to the four novel words (compare to the manipulation in chapter 2). One of the four object-label associations was presented five times (5 Exposure item), one was presented 2 times (2 Exposure item), one was presented once (1 Exposure item), and one object was never presented during the Exposure Phase (0 Exposure item). This manipulation thus created a frequency distribution followed the logic of a small-scale power law distribution, with some objects (e.g. the 5 Exposure item) occurring far more frequently than other objects (the 0 and 1 Exposure item). The resulting eight exposure trials were presented in random order.

In the Yoked Passive condition, participants experienced exactly the same exposure trials in the same order as their yoked counterparts in the Active condition. Participants in the Yoked Passive Exposure Mismatch condition were assigned the same object-label associations in each block as their yoked counterpart; however they experienced a different randomized frequency distribution than their yoked counterparts in the Active condition (Figure 3.2). Note that since these participants experienced a different exposure frequency distribution at random, the degree to which their exposure experience differed from the exposure of their yoked counterparts' experience varied across the sample.

` **Sampling Phase.** After the Exposure Phase, children were given additional exposure to the object-label associations. On a given trial, two objects would appear on the screen at two of the four object locations. To maintain consistency with the Exposure Phase, the objects appeared at the same location that they had appeared in the Exposure Phase. Children were then instructed on how to initiate the next learning trial. In the Active condition, children were instructed to select one of the two objects to appear in the next labeling event (instruction: "Now, you get to choose which friend to learn about. Pick one of Teddy's friends to learn about them."). Once the child tapped on one of the two objects, that object would participate in a labeling event with the same procedure as the labeling events in the Exposure Phase, with the object moving to the center square (while the remaining object locations were obscured) and an auditory stimulus subsequently playing that labeled the selected object (Figure 3.1).

In the practice block of the Active condition, participants completed two sampling trials each involving two of the familiar animal images. In the experimental blocks, participants completed six sampling trials, one for each possible pairing of the four novel objects presented in random order. Thus, participants made a sampling decision between each possible combination of two of the four novel objects (the 0 Exposure item vs. the 5 Exposure item, the 0 Exposure item vs. the 2 Exposure item, etc.).

In the sampling trials of the Yoked Passive and the Yoked Passive Mismatched Exposure conditions, children were instructed to touch the red square to initiate the next labeling event (instruction: "Now, you get to see which friend you will learn about. Touch the square to learn about one of Teddy's friends."). Once the child had triggered the trial by pressing the red square, they next saw a labeling event presented in an analogous procedure to the event from the Exposure Phase. Which object the child saw in these conditions depended on the sampling decision of the Active condition participant that they were yoked to. Each child in the yoked passive conditions experienced the object-label association selected by their yoked counterpart. In other words, children in the yoked passive conditions experienced exactly the same labeling events during the Sampling Phase as a (yoked counterpart) child from the Active condition, with the only difference that the children in the yoked passive conditions did not select their sample (Figure 3.2).

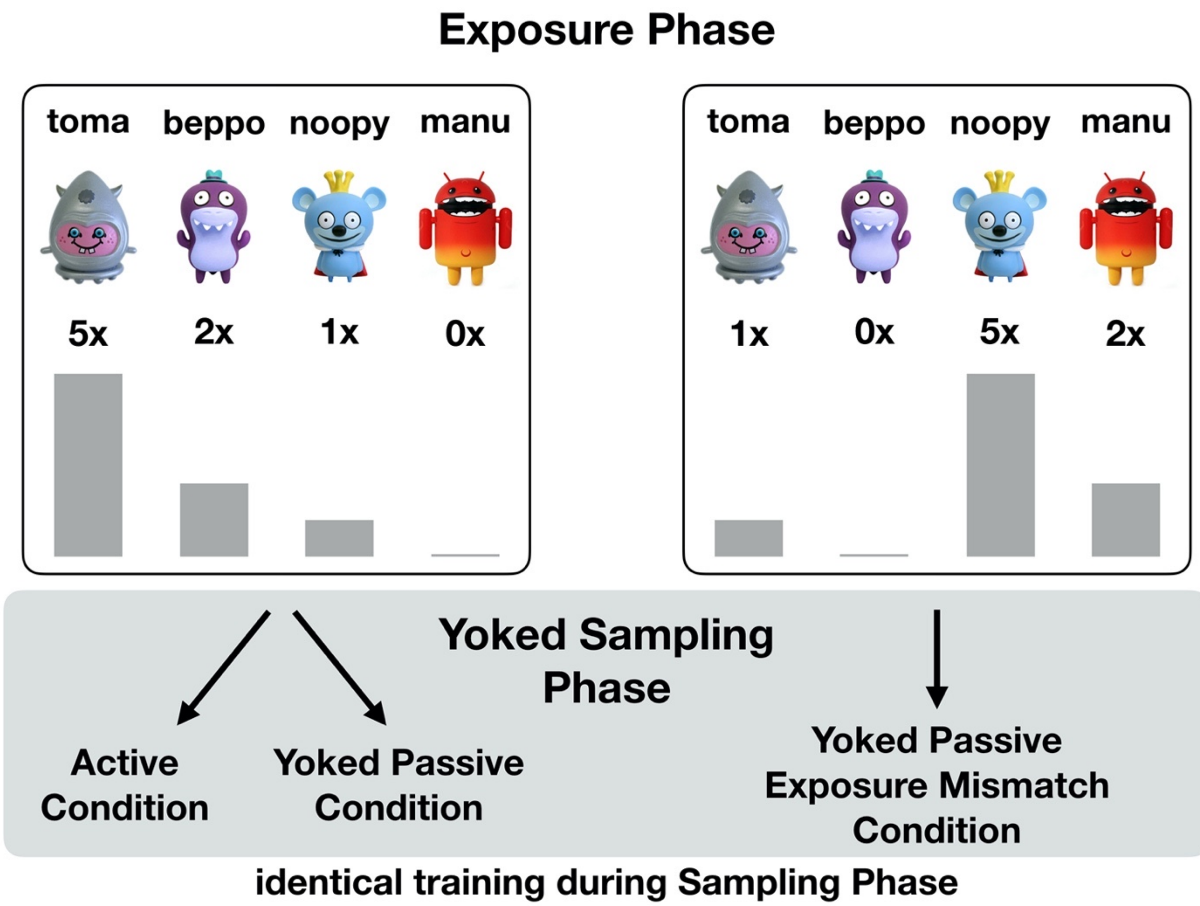

Figure 3.2. Design of the Active, Yoked Passive, and Yoked Passive Exposure Mismatch Conditions.

Test Phase. After completing the sampling phase within a given block, children completed a test phase composed of 2-AFC recognition trials. On each trial, the cartoon bear character would appear at the center of the screen with two objects on either side. Participants were instructed to "tap Teddy to see who she is looking for". Once children touched the cartoon bear character, an auditory stimulus played identifying the target referent, using one of two carrier phrases (*Where's the (target label)* or *Find the (target label)*). Children then had the opportunity to select the object that matched the target label. If children hesitated to respond, the experimenter would prompt children with the target label again with the instruction "Can you (help Teddy) find the (target label)?". The trial ended once children selected one of the two objects. No feedback was provided to children about their accuracy.

In the practice block, children completed two test trials, each involving two of the familiar animal images. In the experimental blocks, children completed eight test trials presented in random order, two for each novel object-label association. The four objects were randomly grouped into two sets of two objects each that always occurred together on a given test trial. This ensured that children could not infer the word meanings through cross-situational learning across the test trials. Each of the objects occurred an equal amount of times at each of the two object locations and occurred as the target object at each location once.

Results

Sampling Patterns – Active Condition

Participants differentially selected words during the Sampling phase depending on the number of occurrences of the word during the Exposure Phase (Figure 3.3). On average, children sampled items that they had never heard the label for during the Exposure Phase far more frequently (M = 1.92, 95% CI = [1.79, 2.05]; maximum number of choices in a given block = 3) than the objects that occurred 5 times during the Exposure phase (M = 1.11, 95% CI = [0.97, 1.25]). Items with an intermediate exposure frequency were sampled at a rate lying roughly halfway between these two extremes in children's sampling preference (2 Exposures: M = 1.45, 95% CI = [1.34, 1.57]; 1 Exposure: M = 1.51, 95% CI = [1.37, 1.65]).

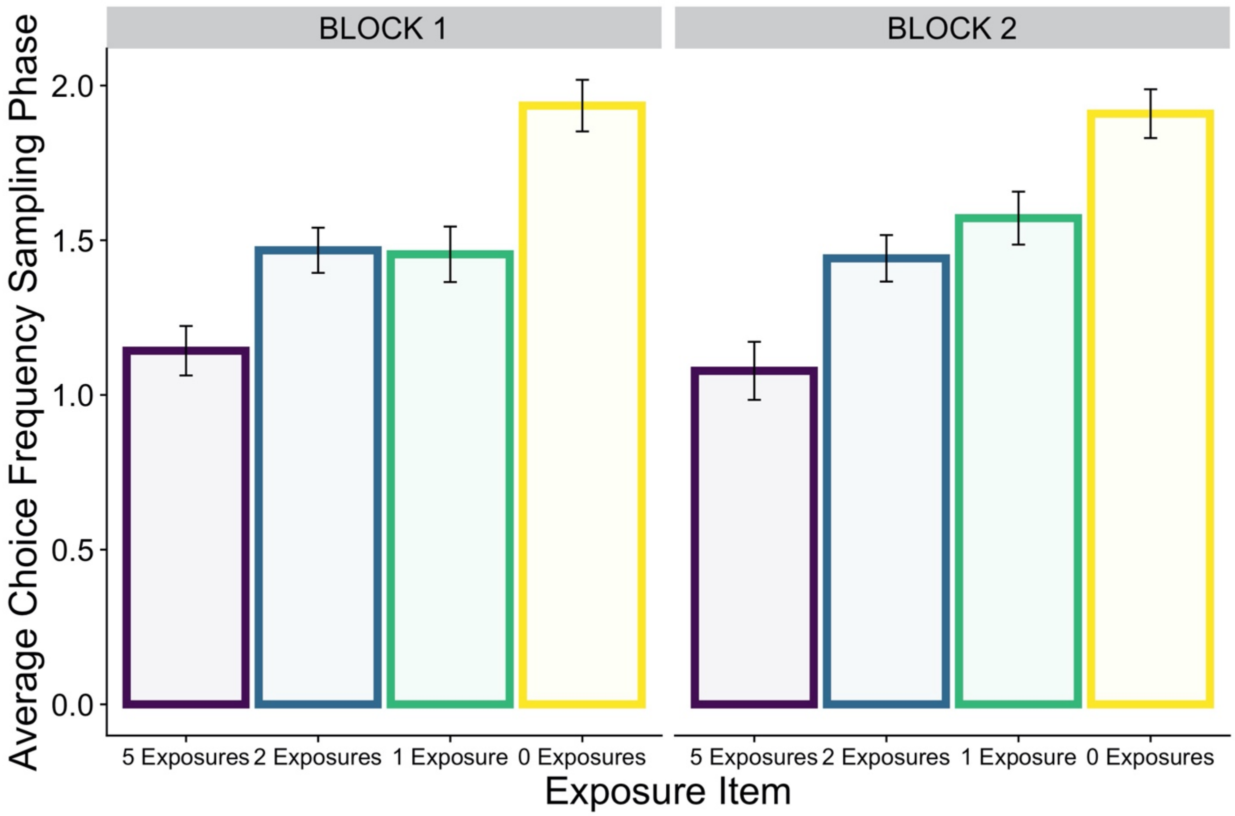

Figure 3.3. Average number of choices dependent on past exposure. The x-axis represents the number of occurrences of the given item during the Exposure phase (split between Block 1 and Block 2). The y-axis depicts the average choice frequency of a given item. Error bars represent +1/-1 SEs.

Next, we asked whether children weighed the relative informativeness of their options in selecting between the two items on a given sampling trial. To investigate this question, we tested whether children preferentially selected the item with the lower frequency during the Exposure phase on a given sampling trial (which always presented two options). Overall, participants in the Active condition preferentially selected words during the Sampling phase that were heard less frequently during the Exposure phase (Block 1: M = 60.8%, 95% CI = [57.0%, 64.7%], $t(76) = 5.57$, $p < .001$; Block 2: M = 61.5%, 95% CI = [56.9%, 66.0%]), $t(76) = 5.01$, $p < .001$). Figure 3.4 depicts the likelihood of choosing the lower exposure item depending on the combination of more or less frequently experienced words during the Exposure phase of blocks 1 and 2.

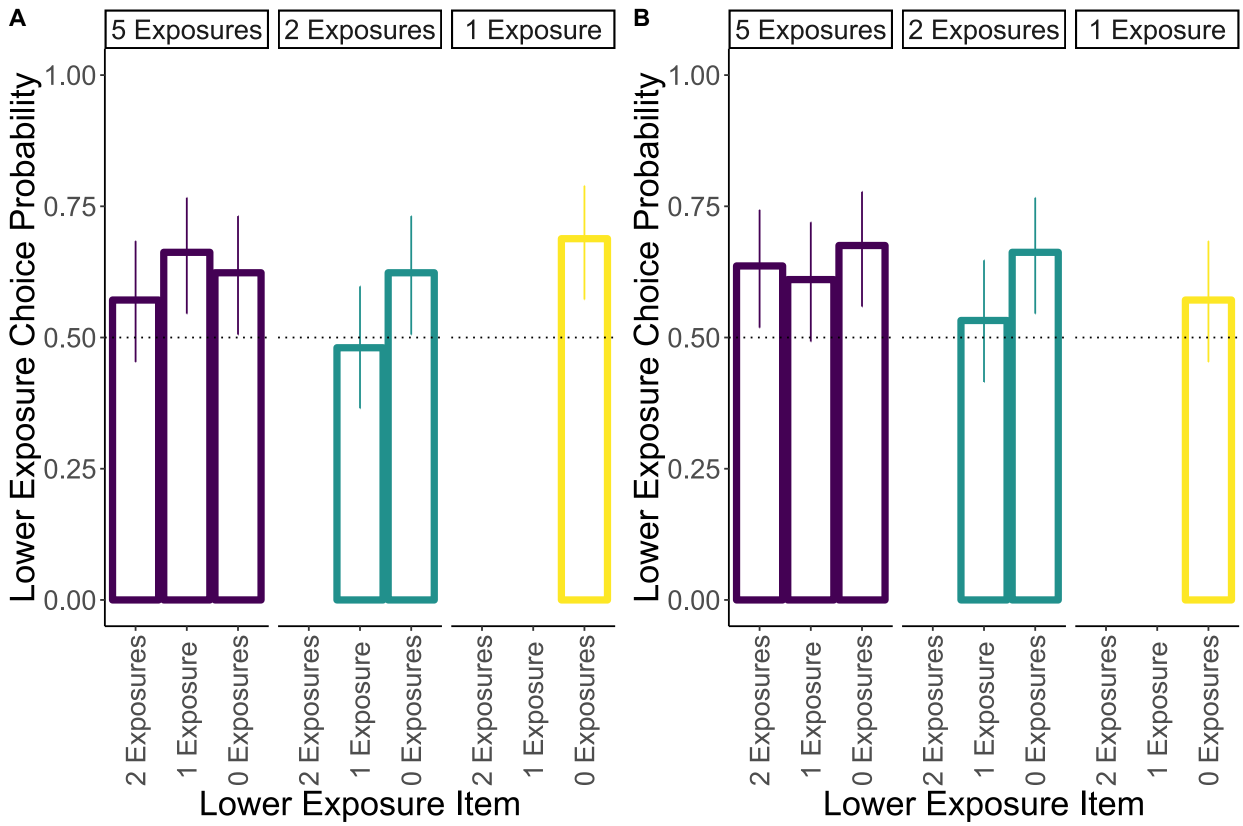

Figure 3.4. Proportion of choices of the lower exposure item during sampling trials on (A) Block 1 and (B) Block 2. The lower exposure option on the given trial is represented on the x-axis, and bars are grouped by the higher exposure item option on the trial (5, 2 or 1 occurrences during the Exposure phase). For example, the first bar from the left depicts the likelihood that children selected the item heard twice during Exposure, given that the alternative is the item heard 5 times during Exposure. Error bars represent 95% CIs.

Informativeness Analysis. To further investigate the patterns in children's sampling behavior across trials, we sought to account for how object-label experience was changing from trial to trial as the sampling phase unfolded. We reasoned that the relative frequency with which children heard a given object-label pair changed with each sampling trial as a consequence of hearing additional object-label input. We therefore conducted a trial-by-trial analysis to investigate whether infants were sensitive to the informativeness of object-label associations based on the entirety of their training experience.

To create a measure of informativeness, we first computed the relative frequency of exposure to a given object-label association *i* as the proportion of trials on which children heard the given object labeled:

$$\text{relative frequency}_i = \frac{\text{trials labeling object i}}{\text{total number of training trials}}$$

For example, on the first trial of the sampling phase (immediately after the exposure phase), the relative frequency of the most frequently experienced object-label pair on this measure was 5/8 = .625, since the object was labeled on 5 of the 8 training trials in the exposure phase. Crucially, this measure takes into account how the relative exposure frequency to each object-label association changes across the sampling phase. For instance, if a participant selected the most frequently seen object-label pair on their first sampling trial, the relative frequency of that particular object-label association would increase to 6/9 = .67. To calculate the corresponding informativeness of the object-label association *i*, we took the standard measure of the negative log relative frequency (see e.g., Barto, Mirolli, & Baldassarre, 2013):

$$\text{informativeness}_i = -\log_2 \text{relative frequency}_i$$

Since $\log_2(0)$ is not defined, we set the relative frequency of the object-label association with a frequency of 0 to be half of the lowest possible relative frequency, assuming a single labeling event (e.g., 1/16 at the completion of the 8 training trials).

To test whether children were sensitive to differences in informativeness between object-label associations, we fit a logistic mixed-effects model predicting children's likelihood of choosing a given item from its difference in informativeness to the alternative option (collapsing across blocks). We included random intercepts for participant and for each of item option. We found that children's likelihood of choosing an item increased as it became more informative relative to the alternative item presented on a given trial, b = 0.41, 95% Wald CI = [0.32, 0.49], *z*

= 9.34, $p < .001$ (Figure 3.5). This effect did not interact with block, and similar effects were observed when each block was considered in isolation.

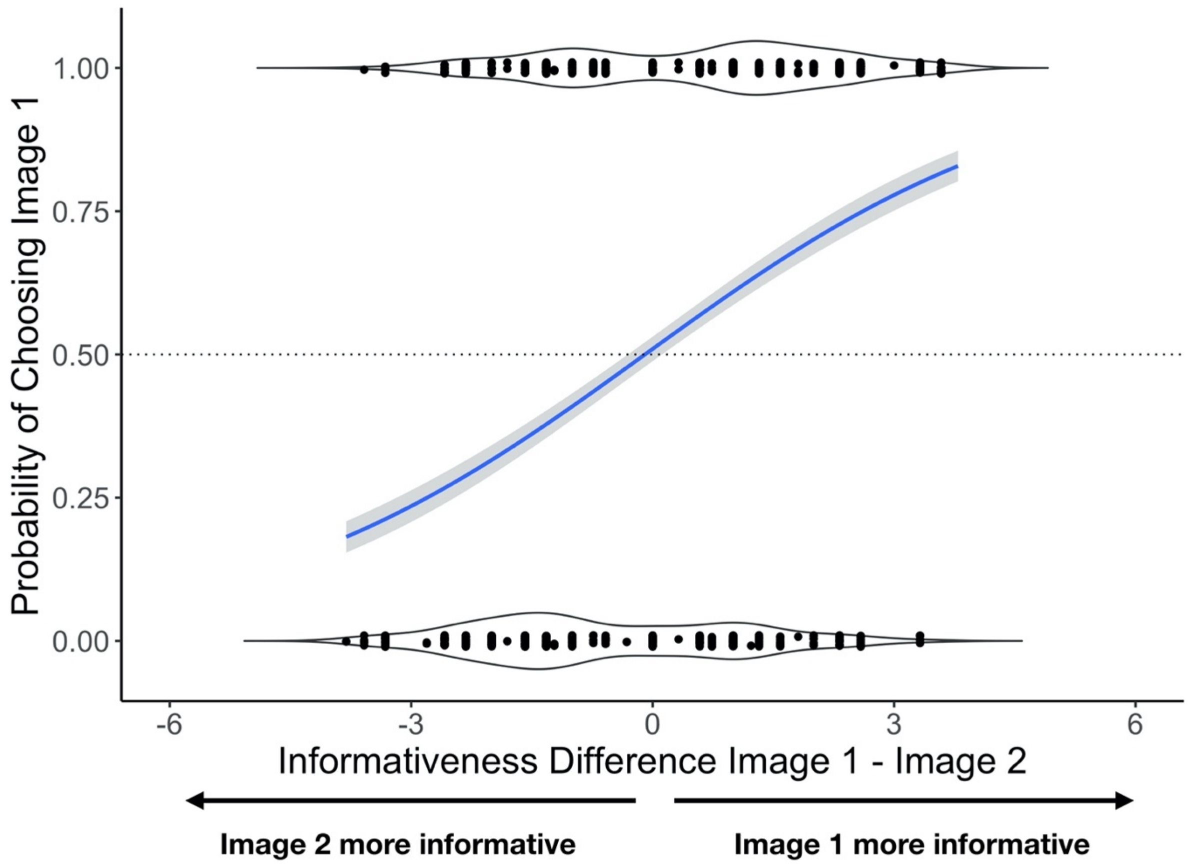

Figure 3.5. Likelihood of choosing an item ("Image 1") as a function of the difference in informativeness to the alternative option ("Image 2") (collapsing across both experiment blocks).

The Effect of Condition on Test Performance

To test for differences in word learning across condition, we fit a logistic mixed-effects model predicting participants' trial-by-trial accuracy from condition (dummy coded) for each experiment block. We included random intercepts for participants, items and for yoked pairings (i.e., observations from participants who were yoked together were treated as non-independent).

In Block 1, we found a marginal overall effect of condition on test accuracy, $\chi^2(2) = 4.82$, $p = .09$. Pairwise comparisons revealed that children performed better in the Active condition (M

= 84.9%, 95% CI = [80.9%, 88.9%]) than in the Yoked Passive Exposure Mismatch condition (M = 78.2%, 95% CI = [73.4%, 83.1%]), b = 0.52, Wald 95% CI = [0.06, 0.99], z = 2.20, p = .028, but not in the Yoked Passive condition (M = 81.8%, 95% CI = [77.5%, 86.1%]), b = 0.27, Wald 95% CI = [-0.20, 0.74], z = 1.13, p = .26 (Figure 3.6).

In Block 2,the overall effect of condition on test accuracy was not significant, $\chi^2(2)$ = 2.69, p = .26. Accuracy did not differ between the Active condition (M = 78.9%, 95% CI = [73.7%, 84.1%]) and the Yoked Passive Exposure Mismatch condition (M = 80.0%, 95% CI = [75.6%, 84.5%]), b = -0.04, Wald 95% CI = [-0.53, 0.46], z = -0.15, p = .88, or between the Active Condition and the Yoked Passive condition (M = 74.2%, 95% CI = [68.8%, 79.6%]), b = 0.33, Wald 95% CI = [-0.16, 0.82], z = 1.33, p = .18.

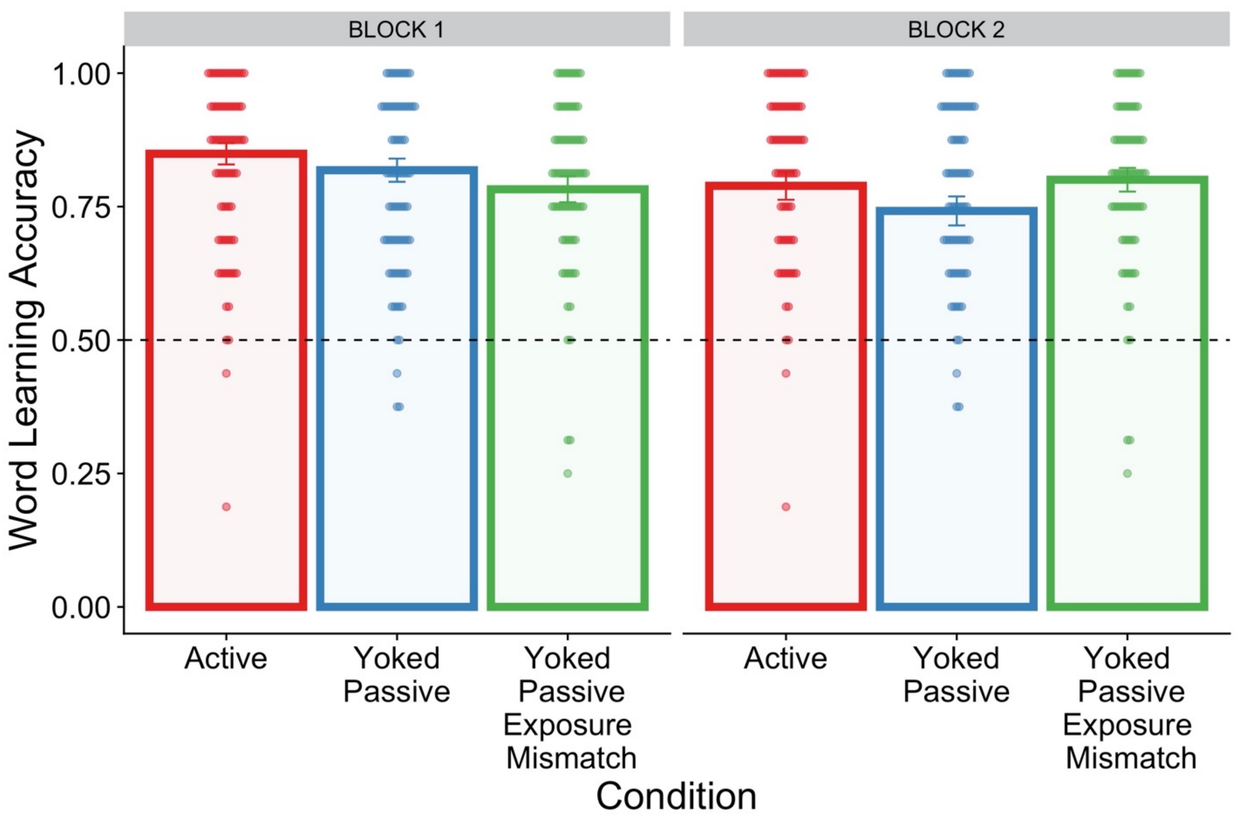

Figure 3.6. Word learning accuracy across conditions. Individual data points are partially overlapping to visualize the distribution of test accuracy. Error bars represent +1/-1 SEs.

Relationships between Exposure, Sampling and Test

We conducted several exploratory analyses to investigate the nature of the relationship between participants' training during the Exposure Phase and the Sampling Phase.

Frequency of experience during both the Exposure Phase and the Sampling Phase predict test accuracy. First, we asked whether the frequency of children's exposure to each object-label association during the Exposure Phase and the Sampling Phase would predict their later test accuracy. We fit a logistic mixed-effects model predicting children's trial-by-trial accuracy on a given item from the frequency with which they experienced that item during the Exposure Phase and the frequency of experiencing that item during the Sampling Phase. We included random intercepts for participants, items and for yoked pairings. Children were more accurate on items they experienced more frequently during the Exposure Phase, $b = 0.08$, Wald 95% CI = [0.03, 0.12], $z = 3.13$, $p = .002$. After controlling for frequency during the Exposure Phase, children were also more accurate on items they experienced during the Sampling Phase, $b = 0.12$, Wald 95% CI = [0.004, 0.23], $z = 2.03$, $p = .04$. The effects of frequency during the Exposure Phase and the Sampling Phase did not interact with block or condition.

Exposure frequency mismatch predicts test accuracy in the Yoked Passive Exposure Mismatch condition. Next, we sought to understand the relationship between the decrease in accuracy in the Yoked Passive Exposure Mismatch condition and the degree of mismatch in exposure frequency for yoked participants in the passive and active conditions. Specifically, did participants in the Yoked Passive Exposure Mismatch condition perform better on items they had more exposure to relative to their yoked counterpart in the Active condition and worse on items they had less exposure to relative to their yoked counterpart?

To investigate this question, we fit a logistic mixed-effects model predicting the trial-by-trial accuracy of participants in the Yoked Passive Exposure Mismatch condition from the difference in exposure frequency for a given item (exposure frequency in the Yoked Passive Exposure Mismatch condition – exposure frequency in the Active condition for a given yoked pair and item). We included random intercepts for participants and items. Accuracy in the Yoked Passive Exposure Mismatch condition increased with larger exposure frequencies relative to the Active condition, $b = 0.06$, Wald 95% CI = [0.002, 0.11], $z = 2.04$, $p = .042$. In other words, when participants in the Yoked Passive Exposure Mismatch condition had a large benefit in exposure frequency relative their counterparts in the Active condition, their accuracy at test tended to be higher. Conversely, when Yoked Passive Exposure Mismatch participants had experienced a given item far less than their yoked counterparts, they tended to have lower accuracy at test. This effect of exposure frequency mismatch did not interact with experiment block. See Appendix B for an additional analysis and visualization (Figure B1) investigating the difference in accuracy between yoked pairings in the Active and Yoked Passive Exposure Mismatch condition, depending on differences in exposure frequency.

Preference for sampling lower exposure items and test accuracy. To investigate the relationship between individual differences in sampling behavior and test performance, we asked whether children who preferentially sampled lower exposure items learned the novel words better. If sampling lower exposure items represents a particularly useful sampling strategy, then preferentially selecting items heard less frequently during training should lead to better test performance. However, there was no significant relationship between preference for selecting lower exposure items and test accuracy, $r = -.01$, $p = .92$.

Discussion

The results from the study support three main conclusions. First, children in the Active condition made choices that matched well with their initial word learning experience: children tended to select those words that they had experienced less frequently during the Exposure phase and showed graded sensitivity in their samples to their previous exposure. Second, children in the Yoked Passive Exposure Mismatch condition, who passively observed samples that were selected by children in the Active condition with differing initial word exposure, showed worse learning than children in the Active condition (at least in Block 1). Third, there was no strong evidence for a difference in word learning between the Active and the Yoked Passive condition – i.e., children who experienced exactly the same word learning experience performed similarly, regardless of whether they were selecting their own input during the Sampling phase.

Why did children in our study not show stronger differences between active and passive learning? The current design ensures that children in the Active condition and the Yoked Passive condition have the same previous learning exposure. Thus, the selections made by children in the Active condition should also benefit children in the Yoked Passive condition, given that children in both conditions had similar uncertainty about the same word meanings. This aspect of the design differs somewhat from past category learning studies that have documented differences between active and yoked passive manipulations (MacDonald & Frank, 2016; Markant & Gureckis, 2014). In these studies, participants are likely to be entertaining differing hypotheses about the category rule at any given point in the experiment, and thus the data generated by one participant testing one hypothesis may not be useful to a yoked participant entertaining a different hypothesis. The consistent manipulation of word learning frequency in children in the Active and Passive Yoked condition means that children in both conditions are likely to be seeking similar information about the novel words, thus reducing the potential benefits of hypothesis-dependent sampling.

For children in the Yoked Passive Exposure Mismatch condition, however, the sample selected by children in the Active condition should not be as beneficial for learning, given their differing experience. This difference in past exposure may underlie the difference in performance between the two conditions in the first block of the experiment, and builds further evidence that children's active sampling decisions are tied to past learning exposure and benefit future learning. Future analyses will seek to systematically quantify the degree of mismatch between children's experience and their input in the Sampling phase, to test whether this mismatch predicts children's test accuracy. Additional support for the notion that children in the Active condition make choices that are well-suited to learning could be built by comparing performance in the Active condition to a receptive condition in which children experience random samples of the novel words (see also Sim et al., 2015). Overall, preliminary results provide evidence that children sample novel words in a manner that is supportive of their word learning.

One outstanding puzzle is why effects differed across the two blocks of the experiment. Originally, we reasoned that having children perform the experiment twice, in two identically structured blocks, would allow children to familiarize themselves with the task and gain a better understanding of the learning goals by the second block of the experiment. Therefore, we expected that, if anything, condition effects would be more pronounced in the second round of the experiment. One possible explanation for the difference in results is declining engagement over the course of the experiment. Accuracy overall was lower in the second block of the experiment, suggesting that children may have been less attentive during the second round, leading to more similar performance across conditions. Considering that benefits of the active learning condition were not predicted to be limited to the second block and central condition comparisons arose opportunistically, further replication is needed to increase confidence in the effect. Future replications could also more specifically test how specific mismatches in previous experience affect how useful a novel sample is for learners, to investigate the degree to which children's sampling decisions are linked to their own prior learning.

Together with the findings with Chapter 2, the present findings show that children selectively sample new words and can construct word learning situations that are beneficial for learning by preferentially selecting informative options. These findings build on our growing knowledge of the benefits of children's active involvement in their own word learning (Begus et al., 2014; Lucca & Wilbourn, 2018b; Oudeyer et al., 2019), demonstrating that children are motivated to generate informative word learning situations. However, we still know little about the developmental trajectory of children's information-seeking skills. Word learning is a particularly daunting task early in development, when there is little existing word knowledge and experience to build from. Are even young infants able to seek new information about words in a targeted fashion? In Chapters 4 and 5, we explore the limits of children's information-seeking skills, using gaze-contingent eye-tracking methods to investigate how systematically infants (under the age of two) sample novel words.

Chapter 4: Do infants systematically sample novel object-label associations?

Are infants passive absorbers of the information around them, or are they active seekers of new information? A long tradition of theories in cognitive development conceptualize children as playing central roles in their own learning process (Piaget, 1955, 1964; Vygotsky, 1978). Piaget, in contrasting his view with behaviorist approaches, emphasizes the importance of children's active engagement with the environment for development: "Learning is possible only when there is active assimilation. It is this activity on the part of the subject which seems to me underplayed in the stimulus-response schema" (Piaget, 1976: p. 77). For example, as they explore their environment, infants will begin to point to objects in their environment to indicate interest and request information (Begus & Southgate, 2012; Kovács et al., 2014; Lucca & Wilbourn, 2018b) and infants' gestures are often met by caregivers labeling or providing other information about the objects that infants are attending to (Wu & Gros-Louis, 2014, 2015).

Yet we still know little about the strategies that infants use to seek out novel information. What kinds of learning events are infants drawn to when given control over their learning environment? One interesting result is that infants are drawn to sequences that are neither too complex nor too simple – sequences that are 'just right' in terms of their complexity (Kidd et al., 2012, 2014). Past research has also shown that infants and young children will seek out new evidence after surprising events (Stahl & Feigenson, 2015) or after events in which they have ambiguous evidence about how a novel object works (Cook, Goodman, & Schulz, 2011; Gweon & Schulz, 2008; Schulz & Bonawitz, 2007). These results suggest that infants can identify and seek out learning events that have the potential to be informative. These findings lay the foundation for a tantalizing explanation for the substantial learning abilities of young children. Children are excellent learners not only due to learning mechanisms, but also due to curiosity-based cognitive mechanisms and motivations that drive children to seek out informative learning events (Kidd & Hayden, 2015; Mani & Ackermann, 2018; Oudeyer et al., 2019; Oudeyer & Smith, 2016).

Curiosity could play an important role in how children learn new words (Mani & Ackermann, 2018). Children show impressive gains in vocabulary between their first and second years of life (P. Bloom, 2000). Explanations have appealed to gains in social learning abilities (Akhtar & Tomasello, 2000), the maturation and accumulated evidence from attention-based mechanisms (Samuelson & Smith, 1998; Yu & Smith, 2012), and their combination (Hollich et al., 2000). An important contributor could also be infants' active role in seeking out new information and constructing their learning input (L. Bloom, 2000b; Hidaka et al., 2017; Mani & Ackermann, 2018; Oudeyer et al., 2019; Smith et al., 2018). Children's curiosity would be a particularly influential driver of word learning if it led infants to seek out learning events that provide new information about words. Do infants make choices that are informative given their past word learning experience?

One prominent reason why some word learning events are more informative than others for infants is that their exposure to words is not distributed evenly – some words occur more frequently than others (Zipf, 1965). Frequency of word occurrence has strong effects on children's word learning (Ambridge, Kidd, Rowland, & Theakston, 2015). Word frequency is one of the best predictors of when a word will enter a child's vocabulary across languages (Braginsky, Yurovsky, Marchman, & Frank, 2019). The frequency with which a word occurs within a caregiver's speech predicts children's later knowledge of that word (Goodman, Dale, & Li, 2008; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991). Moreover, the frequency with which children are exposed to novel-object label associations in experimental settings predicts their success at learning these new words (Vouloumanos & Werker, 2009). Thus, words that children have heard less frequently – and therefore are less likely to have learned successfully – are more informative for young learners.

In the current work, we asked whether 20-month-old infants differentially sample information about novel word-object associations when given active control over what material they can learn about. We first exposed infants to object-label pairs in a skewed distribution: while all objects occurred equally often, some were labeled more frequently than others. We then used a novel eye-tracking methodology in which infants could control which object-label pairings they would hear next. Our hypothesis was that infants would preferentially sample object-label associations that were more informative based on their past word learning experience.

Experiment 1

In Experiment 1, we tested whether 20-month-old infants would be sensitive to the informativeness of object-label choices when given the ability to initiate learning trials in a gaze-contingent eye-tracking design. To manipulate the informativity of novel object-label associations, we first provided infants with initial exposure to four novel object-label associations during a training phase. Half of the objects were labeled on all four training trials (high-frequency object-label associations), while the other half of objects were presented with a label on only one of their four training trial occurrences (low-frequency object-label associations). Infants were then given the opportunity to sample additional label information for the four objects. The central question was whether infants would be sensitive to the informativity of different labeling events based on their past word learning exposure.

We targeted infants in the second half of their second year for two reasons. First, infants begin to vastly improve their ability to flexibly shift and sustain attention to individual objects and in the presence of multiple objects between 1 and 2 years of age (Colombo, 2001; Ruff & Capozzoli, 2003). Second, as infants approach their second year of life, observational and corpus-based studies suggest that children show a high proportion of queries that ask for labels (e.g., "What is that?") or explanations of the reference of novel words ("What's a jack-o-lantern?") (Chouinard, 2007). Thus, infants in this age range are likely to have the necessary visual attention skills to control gaze-contingent events and may be developing an interest in seeking information about novel labels.

Methods

Participants

In study 1, twenty-five 19- to 21-month-old infants (10 female) participated in the experiment (mean age = 20.2 months, SD = .72). An additional 12 infants were tested but excluded from the analyses due to fussiness (n = 8), experimenter error (n = 2) or demonstrating a strong side bias during testing (defined as looking to a single side on over 75% of sampling trials; n = 2). All infants were English-learning, full-term infants with no vision or hearing problems and no exposure to a second language. Families were recruited from the Madison, WI community and received a gift for participating in the study.

Stimuli

The novel objects consisted of four unfamiliar objects from the Novel Object and Unusual Name Database (see Figure 2.1; Horst & Hout, 2016) and four novel labels (*tursa*, *gasser*, *permi*, *labo*). Each object was selected to be distinct in shape and color. All pictures of objects were presented against a white background. We created four object-label pairings that were divided into two sets of two object-label pairings. The two sets were randomly assigned to be in the high-frequency and low-frequency labeling condition during training, and the assignment was counterbalanced across infants.

All auditory stimuli were spoken by a female native English speaker from Wisconsin in child-directed speech. The novel words were normalized for intensity and duration such that the duration of all novel words was equal. The words were presented in carrier phrases that were normalized for intensity.

Apparatus and Stimuli Presentation

Participants were tested using a Tobii T60 XL eye-tracker with a sampling rate of 60 Hz and displayed on a 20.5'' screen. The auditory and visual stimuli were presented in PsychoPy (Peirce, 2007). Communication with the eye-tracker and gaze-contingent stimulus presentation was controlled using the PyGaze toolbox (Dalmaijer, Mathôt, & Van der Stigchel, 2014). The infants sat on their caregivers' laps approximately 60 cm from the monitor in a sound-attenuated booth. To prevent caregivers from influencing infants' behavior, caregivers wore darkened glasses. A five-point calibration sequence was used to obtain a reliable track of participants' looking location (four corners plus center).

Design & Procedure

Training Phase. At the beginning of the training phase, all four objects were presented simultaneously on the screen along with an attention-getting music recording for 10 s. Infants then entered a labeling phase in which they heard novel labels paired with novel objects. On each trial, an object appeared on the right or left corner of the screen in silence for 1000 ms. Then, the object moved vertically up and down on the screen for 4000 ms. The object was accompanied by speech that either labeled the object using one of two carrier phrases (*Look at the tursa. That's a* *tursa*. or *See the tursa. It's a tursa*.) or directed infants' attention to the object without providing a label (*Wow! Look at that! Do you like it?*). Then, the object remained on the screen for an additional 1000 ms in silence before the trial ended and the next trial began after an inter-trial interval of 500 – 700 ms (jittered). On the first two labeling trials, infants were presented with two familiar objects (a ball and a shoe) and their labels, to familiarize them with the task. Next, participants were presented with four blocks in which each of the four objects were presented once in random order, for a total of 16 trials. Each block of four trials was followed by a short attention-getting trial in which a nature scene was presented for 5000 ms together with attention-getting phrases to maintain infants' attention and reduce the repetitiveness of the task.

The key manipulation was the frequency with which each object occurred with its label (Figure 4.1). For each infant, two objects were labeled on all four training trials (*high-frequency* object-label associations). The other objects were only labeled on the first training trial (*low-frequency* object-label associations) and were presented with attention-getting audio without label information on the 3 subsequent trials. Consequently, while all objects were seen equally often, two objects were labeled four times more frequently than the other two objects. The frequency role of the object pairs (high vs. low frequency) was counterbalanced across infants.

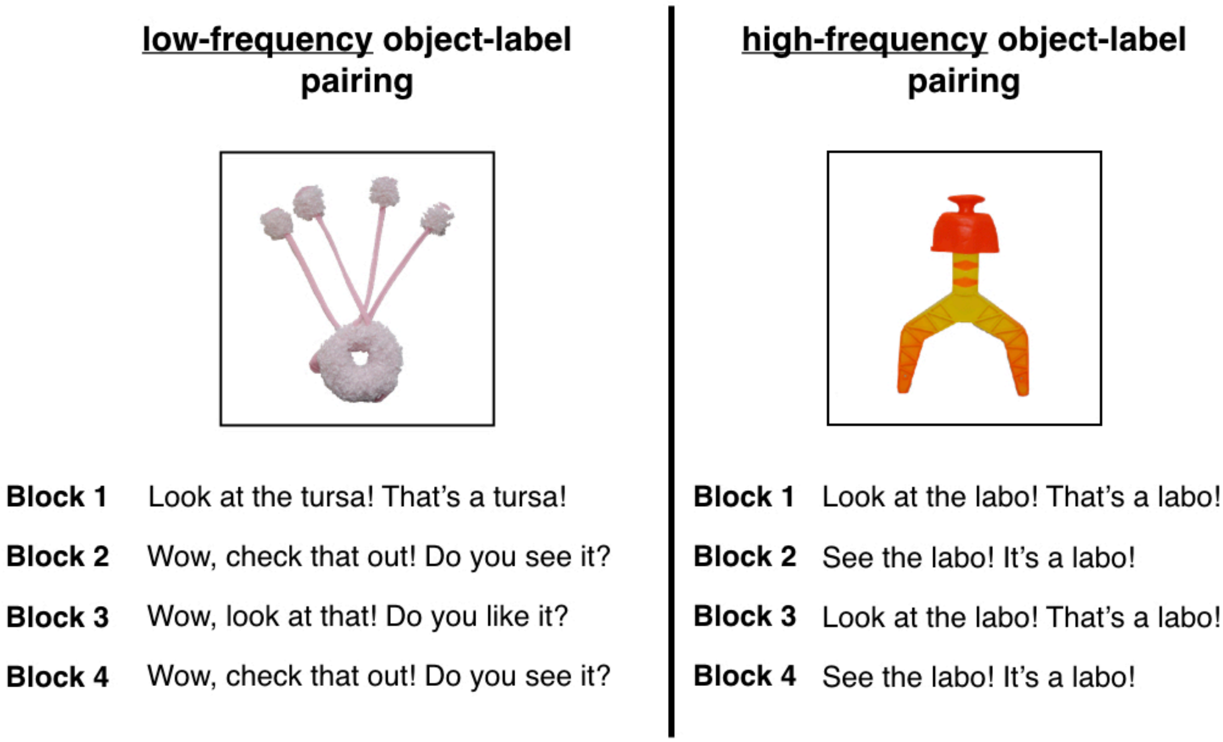

Figure 4.1. Manipulation of frequency of object-label pairings in the Training Phase.

Sampling Phase. In the subsequent *Sampling Phase*, infants controlled which object-label association they heard next through their gaze fixations (Figure 4.2). On each trial, two objects appeared on the screen. One of the objects was always a high-frequency object-label pair from the training phase, and the other was always a low-frequency object-label pair (Figure 4.1). After the objects were shown in a stationary position for 1000 ms, we drew infants' attention to each object by having each object "wiggle" from left to right for 500 ms in turn, with 2000 ms of static exposure between the two "wiggles". The order in which infants' attention was drawn to each object was counterbalanced across infants. The objects were then presented for an additional 1000 ms in a static position, after which two screens slid down to cover the two objects for 500 ms. This initiated the gaze-contingent portion of the trial, during which infants' fixations controlled which object was labeled. We obscured the two objects to minimize the extent to which infants' fixations were driven by perceptual properties of the objects. When the duration with which infants fixated one of the two screens exceeded 700 ms, the fixated screen lifted over the course of 500 ms, revealing the object behind the screen. This object was then labeled using the same carrier phrase for all labels (*It's a [label]*) for 1350 ms, followed by 650 ms of silence. If infants' fixations to one of the two screens did not exceed the 700 ms threshold within 5000 ms, one of the screens rose at random and the object was labeled. We excluded trials in which the 5000ms threshold was exceeded and the labeling event was not triggered by infants' fixations from all subsequent analyses (12.5% of all trials).

The sampling phase began with two trials with familiar objects and labels (the shoe and ball from the Training phase) to train participants on the sampling procedure. Infants then received a series of sampling trials using the novel object-label pairs from the training phase. On each trial, an object that had been labeled frequently and an object that had been labeled infrequently during training appeared together. The objects were presented in one of two pseudo-random orders in which object location was counterbalanced. The first 15 infants received 16 sampling trials. This number was halved to 8 sampling trials for the final 10 infants because this longer sampling procedure proved fatiguing for many infants. No difference was found between infants based on number of sampling trials seen in our main analyses, so we present the results together (see Results).

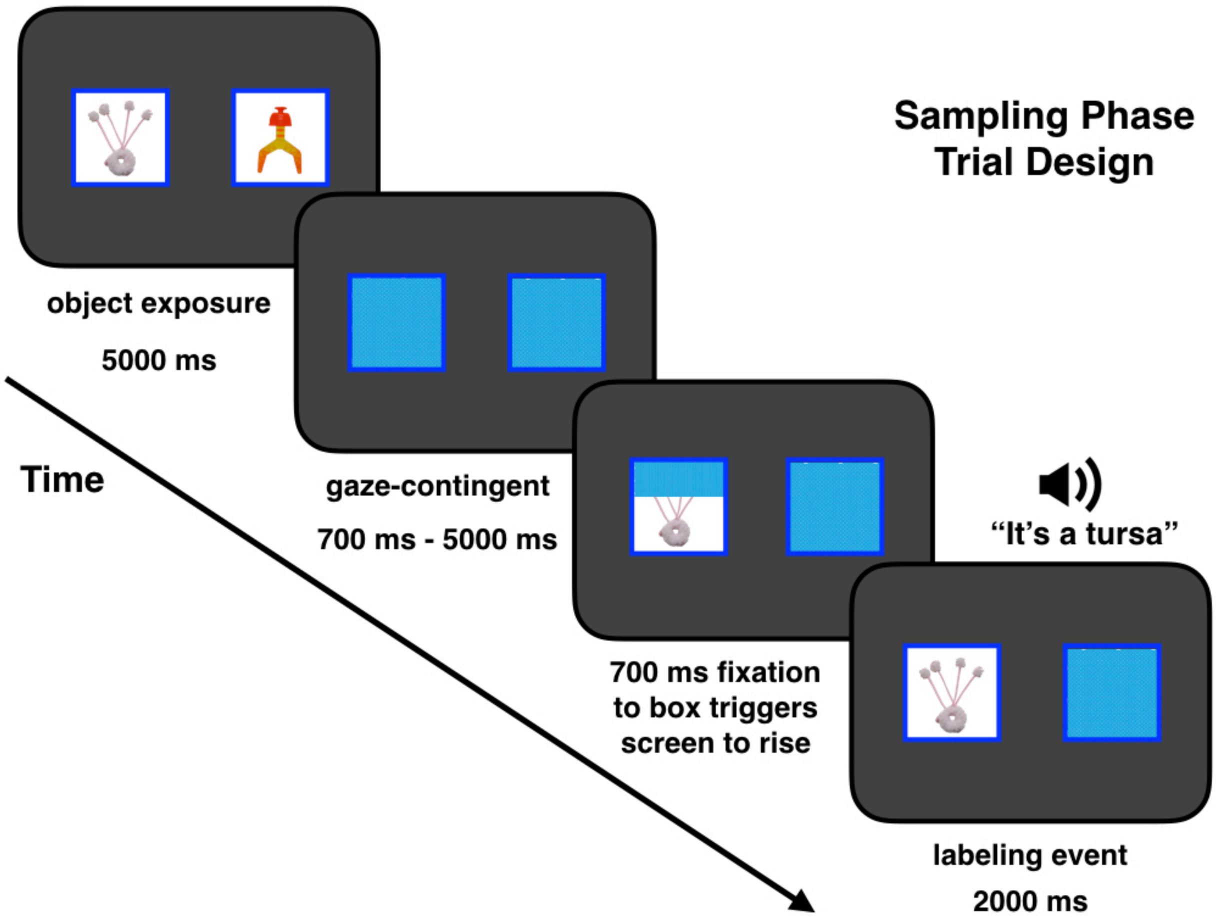

Figure 4.2. Gaze-contingent sampling phase design.

Picture Pointing Test. A subset of the infants (n = 14) also completed a short picture pointing test designed to probe infants' knowledge of the object-label mappings. One additional infant was tested but did not contribute data because they did not make any pointing gestures during the procedure. Infants sat on parents' laps while the experimenter presented them with 4 2-AFC test trials on laminated sheets of paper depicting two objects from training. Each object occurred once as a target and once as a foil with a randomly selected object. The four trials were presented in one of two pseudorandomized orders. On each trial, the experimenter asked the infant "Where's the [label]? Can you point to it?" The experimenter recorded which object the infant pointed to first. The data recorded by the experimenter during the session was subsequently matched to the particular object-label mapping the infant heard in the experiment to determine correctness of the infant's point.

Results

Overall Sampling Preferences

On the gaze-contingent sampling trials, infants were equally likely to choose objects that were presented as high-frequency and low-frequency object-label associations (percent low-frequency choices = 48.6%, 95% CI = [44.2%, 53.1%]). To test whether infants' choices changed over the course of the sampling phase, we fit a logistic mixed-effects model predicting likelihood of choosing the low-frequency object-label association from trial number, including a by-participant random intercept and a by-participant random slope for trial number. There was a marginal decrease in the likelihood of choosing the low-frequency object-label association across trials, $b = -0.05$, 95% Wald CI = [-0.10, 0.01], $z = -1.77$, $p = .077$. Infants were more likely than chance to choose the low-frequency object in the first four sampling trials (percent low-frequency choices = 56.3%, 95% CI = [50.1%, 62.6%], $t(24) = 2.10$, $p = .046$), but became less likely to choose the low-frequency object-label associations as the sampling trials unfolded (percent low-frequency choices in Block 2: M = 44.0%, 95% CI = [33.9%, 54.1%]; Block 3: M = 50.0%, 95% CI = [37.3%, 62.7%]; Block 4: M = 33.3%, 95% CI = [15.5%, 51.1%]). This finding suggests that infants show a weak initial preference for more informative object-label associations. This initial preference declines overall across sampling trials, perhaps due to the memory of the frequently experienced object-label associations decaying across sampling trials (see e.g., Pelz, Piantadosi, & Kidd, 2015).

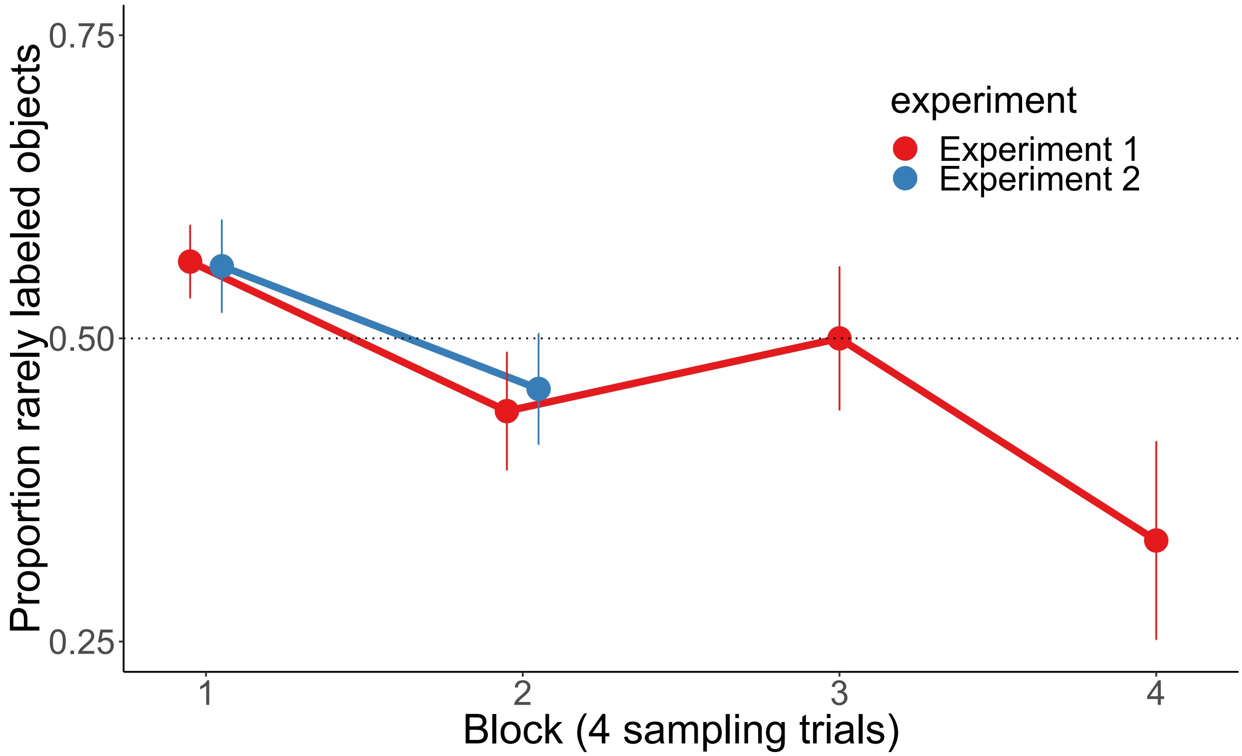

Figure 4.3. Proportion high-frequency object-label associations selected across the Sampling phase in Experiments 1 and 2. The Sampling phase was shortened in Experiment 2 to only include 8 sampling trials. Error bars represent +1/-1 SEs.

Informativeness Analysis

To further investigate the patterns in infants' sampling behavior across trials, we sought to account for how infants' object-label experience was changing from trial to trial as the sampling phase unfolded. We reasoned that the relative frequency with which infants heard a given object-label pair changed with each sampling trial as a consequence of hearing additional object-label input. We therefore conducted a trial-by-trial analysis to investigate whether infants were sensitive to the informativeness of object-label associations based on the entirety of their training experience.

To create a measure of informativeness, we first computed the relative frequency of exposure to a given object-label association *i* as the proportion of trials on which infants heard the given object labeled (cf. chapter 3):

$$\text{relative frequency}_i = \frac{\text{trials labeling object i}}{\text{total number of training trials}}$$

For example, on the first trial of the sampling phase (immediately after the training phase), the relative frequency of the high-frequency object-label pairs on this measure was 4/16 = .25, since the object was labeled on 4 of the 16 training trials. The relative frequency of the low-frequency object-label associations after the Training Phase was 1/16 = .0625. Crucially, this measure takes into account how the relative exposure frequency to each object-label association changes across the sampling phase. For instance, if the infant selected the low-frequency object on their first sampling trial, the relative frequency of that particular object-label association would increase to 2/17 = .12. To calculate the corresponding informativeness of the object-label association *i*, we took the standard measure of the negative log relative frequency (see e.g., Barto, Mirolli, & Baldassarre, 2013):

$$\text{informativeness}_i = -\log_2 \text{relative frequency}_i$$

Entering the sampling phase, the low-frequency object-label associations were twice as informative on this measure (informativeness = 4) as the high-frequency object-label associations (informativeness = 2). Trials in which the informativeness difference between the objects equaled zero were removed, since $\log_2(0)$ is not defined (6 out of 279 total trials). Alternative choices (e.g., setting the relative frequency to half of the lowest observed relative frequency, as in Chapter 3) yield similar results.

To test whether infants were sensitive to differences in informativeness between object-label associations, we fit a logistic mixed-effects model predicting infants' likelihood of choosing the more informative object from the difference in informativeness between their two options using the lme4 package in R (Bates & Maechler, 2009). We included a by-participant random intercept and a by-participant random slope for the difference in informativeness. We found that as the difference in informativeness between infants' options increased, they became more likely to choose the more informative object-label association, b = 0.83, 95% Wald CI = [0.29, 1.36], $z = 3.03$, $p = .002$ (Figure 4.4). The same result holds separately both for the infants who saw 16 sampling trials ($z = 2.10$, $p = .036$) and for the infants who saw 8 sampling trials ($z = 3.63$, $p < .001$). Infants were sensitive to the informativeness of their options when controlling their own learning events: as the difference in informativeness rose between their two word choices, infants were more likely to trigger the more informative object-label association.

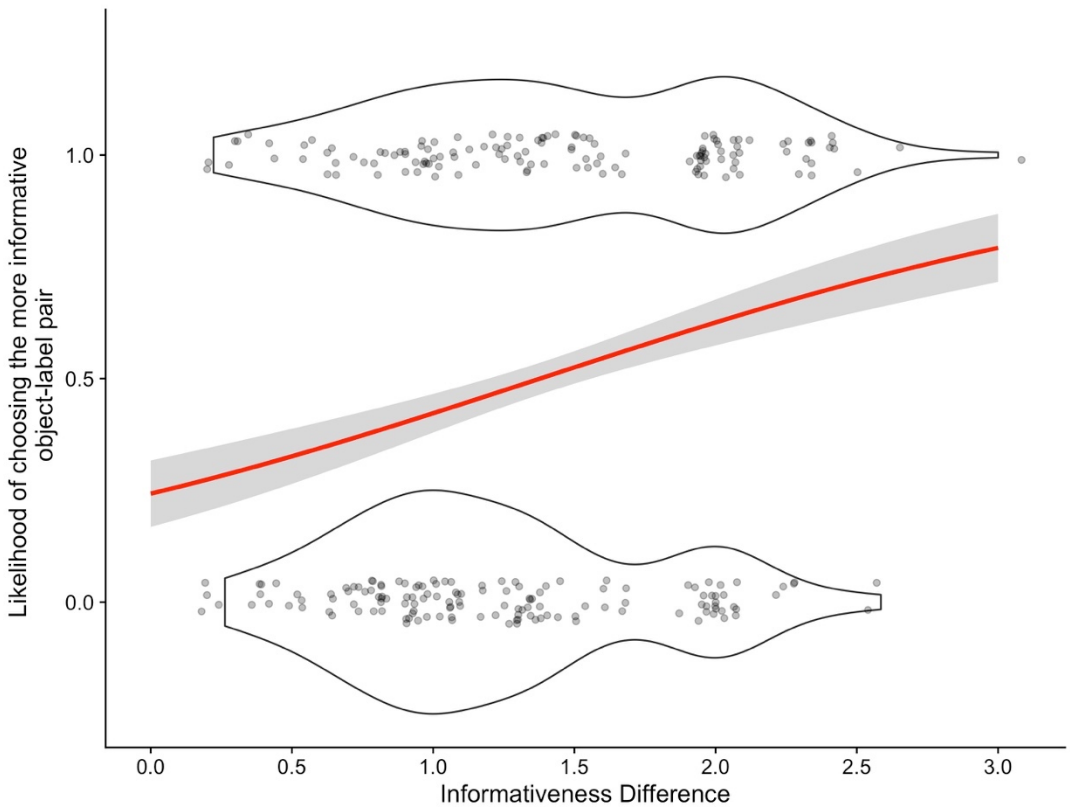

Figure 4.4. Informative Analysis Results in Experiment 1. The difference in informativeness between object-label choices predicts the likelihood of choosing the more informative object-label association. Error bands represent +1/-1 SEs.

Picture Pointing Test Results

Overall, infants showed evidence of learning the object-label associations, M = 72.6%, 95% CI = [57.1%, 88.1%], $t(13) = 3.15$, $p = .008$ (Figure 4.5A). Accuracy was similar for low-frequency (M = 69%, 95% CI = [49.6%, 88.9%]) and high-frequency object-label associations (M = 78.6%, 95% CI = [59.9%, 97.2%]), $t(12) = -1.30$, $p = .22$. To test whether accuracy was predicted by infants' sampling decisions, we fit a logistic mixed-effects model predicting infants' accuracy on each test trial from the frequency with which they chose the target object during the sampling phase. We included a by-participant random intercept and a by-participant random slope. Object-label associations that infants sampled more frequently during the Sampling Phase were marginally more likely to be identified correctly during the picture pointing test, $b = 0.55$, 95% Wald CI = [-.01, 1.11], $z = 1.94$, $p = .053$ (Figure 4.5B).

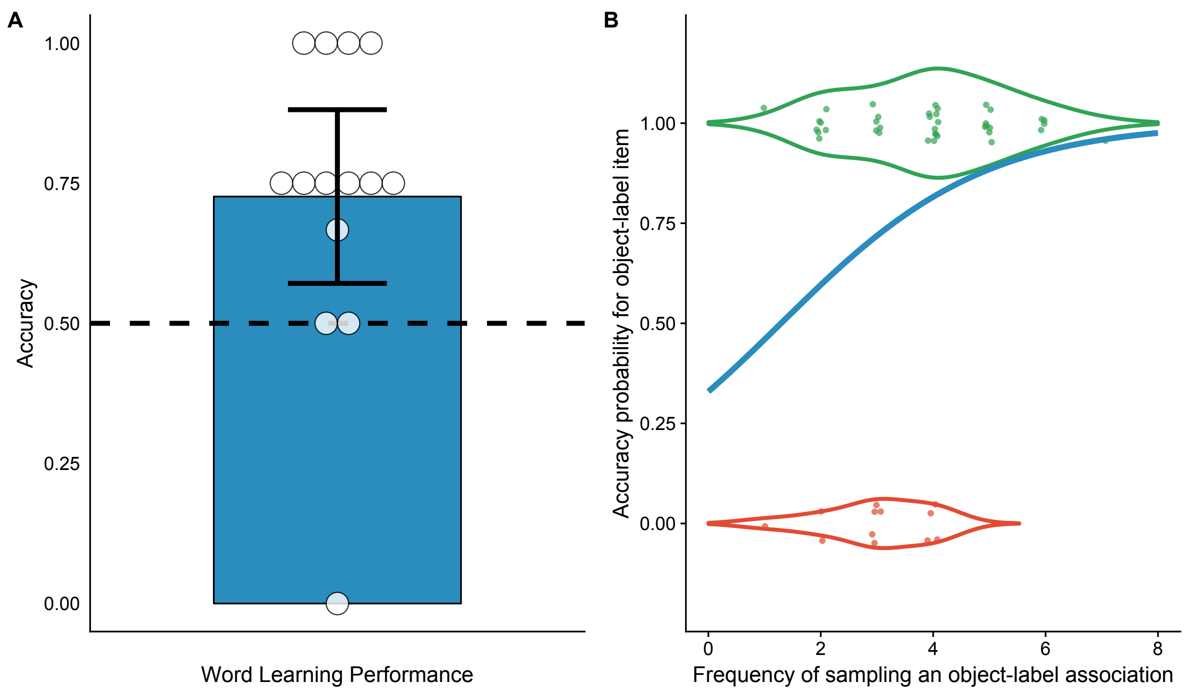

Figure 4.5. Test Performance in Experiment 1. (A) Overall picture pointing test accuracy in Experiment 1 and (B) Frequency of selecting an object-label association during the Sampling Phase relates to test accuracy. Error bar represents 95% CI.

Experiment 2

The goal of Experiment 2 was to replicate the results from Experiment 1, while investigating participants' learning of the novel object-label associations with a looking-while-listening design.

Methods

Participants

In study 2, twenty-eight 19- to 21-month-old infants (15 female) participated in the experiment (mean age = 19.8 months, SD = .64). An additional 9 infants were tested but excluded due to fussiness (n = 6), experimenter error (n = 2), or equipment failure (n = 1).

Design & Procedure

The general procedure was similar to Experiment 1, with three exceptions. First, we included additional sampling trials with familiar objects and labels at the beginning of the experiment to give infants more training with the gaze-contingent procedure. Second, we made a slight modification to the timing of the sampling phase in order to give infants more time to process the images before the gaze-contingent portion of the trial. Third, we included looking-while-listening trials to test word learning after the sampling phase.

Sampling Training. In order to increase the amount of training infants received with the gaze-contingent procedure, infants completed four sampling trials at the outset of the experiment with familiar objects and labels (*ball*, *cat*, *cookie*, *shoe*). The timing of the trials was identical to all other sampling trials in Experiment 2 (see below). The four familiar-word sampling trials were followed by the object-label training phase.

Training Phase. The training phase was identical to Experiment 1.

Sampling Phase. The sampling phase was identical to Experiment 1, with the following modifications. First, infants viewed only one sampling trial with familiar objects (*ball*, *shoe*) since infants were trained on the sampling procedure at the beginning of the experiment. This modification also helped to shorten the gap between the training phase and the first novel object-label association sampling trial. Second, we doubled the amount of time that each object "wiggled" during the initial object exposure, from 500 ms to 1000 ms, allowing infants more time to fixate each object prior to the gaze-contingent portion of the trial.

Looking-While-Listening Test Trials. After the sampling phase, infants completed 2 looking-while-listening trials with familiar objects and labels (*cookie*, *shoe*, *cat*, *duck*) followed by 8 test trials designed to probe infants' knowledge of the novel object-label associations. On each trial, two object images were presented side by side in the center of the screen. After 2000 ms of silence, the test audio began. The test audio consisted of one of two carrier phrases ("Where's the ____?" or "Find the ____."), the target label and a final attention-getting phrase (e.g., "Check that out!" or "Can you see it?") and played for a total of 4000 ms. The 8 novel test trials were presented in two blocks of four trials with a short attention-getting video presented between blocks. Each of the four target labels was tested once in each block. The high-frequency and the low-frequency object-label associations were yoked such that the low-frequency objects always appeared together and the high-frequency objects always appeared together on a given trial as target and foil. The test trials were presented in one of two pseudorandom orders, which were counterbalanced across participants.

Processing and Analyzing Eye-tracking Test Trial Data

Eye movements were recorded and segments of lost data up to 150 ms during which the gaze location did not change were interpolated (S V Wass, Smith, & Johnson, 2013). Areas of interest were set around the two object image positions with an additional 50 pixels around each object to account for imprecision in measurement.

To analyze the looking-while-listening test results, we set a 3000 ms target window stretching from 300 ms post-label onset until 3300 ms post-label onset. Trials on which there was missing data for over 50% of the target window (35.6% of all trials test trials) were excluded. After excluding these trials, we next excluded infants who contributed fewer than half of the total number of test trials (fewer than 4 test trials) from the final analysis (n = 8).

Results

Overall Sampling Preferences based on Training Phase

As in Experiment 1, infants chose high-frequency and low-frequency object-label associations with similar frequency overall during the sampling phase (percent low-frequency choices = 51.4%, 95% CI = [46.4%, 56.4%]). There was little evidence for a decrease in the likelihood of choosing the low-frequency object-label association across trials, $b = -0.08$,95% Wald CI = [-0.20, 0.04], $z = -1.32$, $p = .19$. While consistent with the trend seen in Experiment 1, infants were not more likely than chance to choose the low-frequency object in the first four sampling trials (percent low-frequency choices = 55.9%, 95% CI = [48.1%, 63.8%], $t(24) = 1.55$, $p = .13$) than in the final four sampling trials (percent low-frequency choices = 45.8%, 95% CI = [36.4%, 55.2%], $t(27) = -0.91$, $p = .37$), $z = -1.40$, $p = .16$ (Figure 4.3).

Informativeness Analysis

To test whether infants were sensitive to the difference in informativeness between their choices, we fit a logistic-mixed effects model predicting the likelihood of choosing the more informative object on a given trial from the difference in informativeness between the two options. As in Experiment 1, infants' probability of choosing the more informative object rose as the magnitude of the difference in informativeness increased, $b = 0.85$, 95% CI = [0.23, 1.47], $z = 2.69$, $p = .007$ (Figure 4.6).

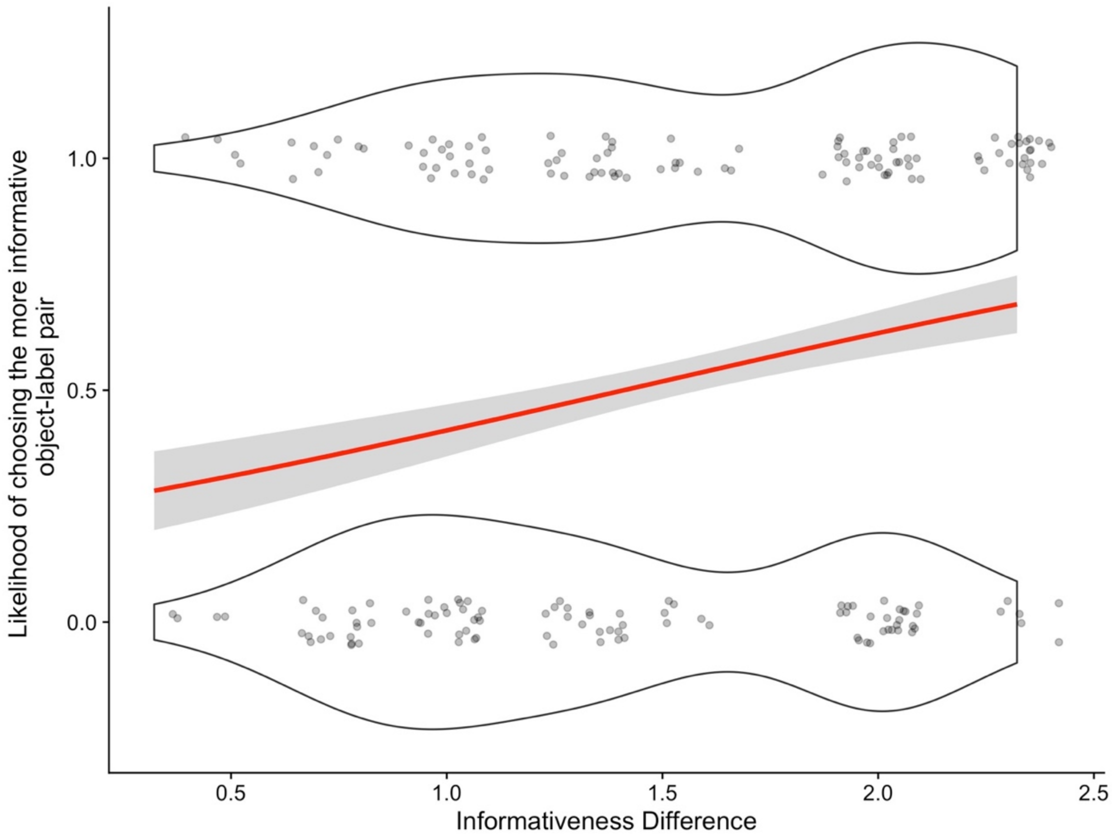

Figure 4.6. Informativeness analysis results in Experiment 2. Error bands represent +1/-1 SEs.

Looking-While-Listening Test

To investigate infants' learning of the novel object-label associations, we analyzed infants' proportion of looking to the target object during the target window. Overall, participants fixated the target object at above-chance levels, M = 60.0%, 95% CI = [53.9%, 66.2%], $t(19) = 3.40$, $p = .003$ (Figure 4.7A). Accuracy did not differ for high-frequency (M = 62.6%, 95% CI = [54.4%, 70.7%]) and low-frequency (M = 57.3%, 95% CI = [48.0%, 66.6%]) targets, $t(19) = 0.92$, $p = .37$. To test whether sampling frequency of the target object-label association during the sampling phase was related to test accuracy, as in Experiment 1, we fit a linear mixed-effect model predicting test accuracy from sampling frequency, including the maximal by-participant random effects structure. Sampling frequency was not related to test accuracy, $b = 0.03$, $\chi^2(1) = .68$, $p = .41$ (Figure 4.7B).

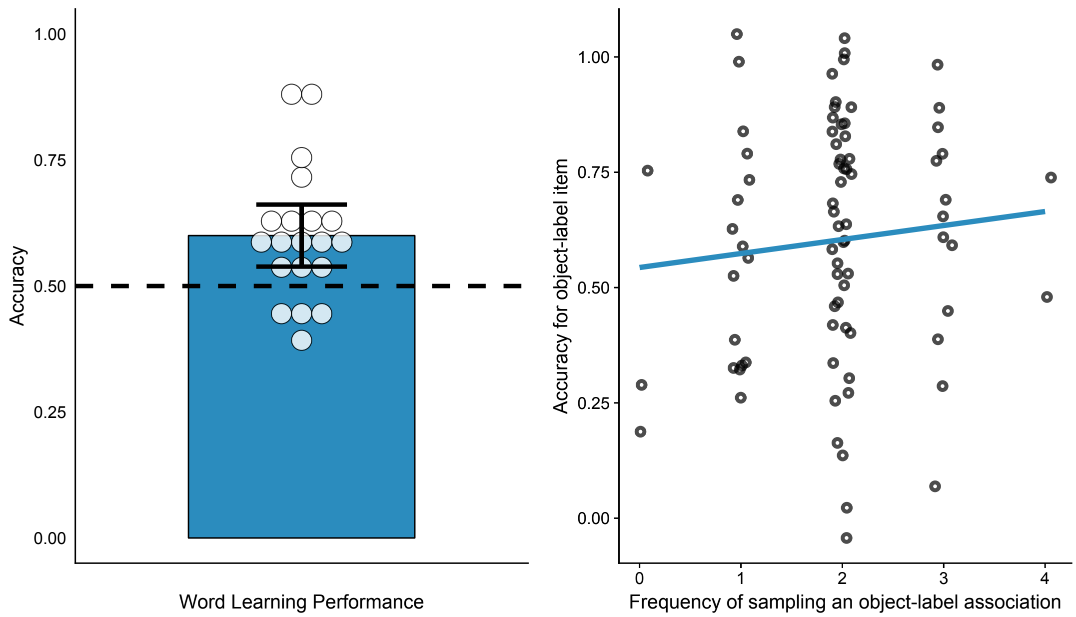

Figure 4.7. Test performance in Experiment 2. (A) Overall looking-while-listening test accuracy in Experiment 2 and (B) no significant relationship between the frequency of selecting an object-label association during the Sampling Phase and test accuracy. Error bar represents 95% CI.

Discussion

In Experiments 1 and 2, we found that infants showed systematic patterns in how they sampled novel object-label associations. Infants showed a small, but consistent, preference for previously less frequently experienced object-label associations at the beginning of the sampling phase. This finding provides preliminary support for the hypothesis that infants show a preference for more informative object-label associations. In analyzing infants sampling patterns across the entirety of the sampling phase, we found that infants showed a broader sensitivity to the relative informativeness of different sampling events. The likelihood of infants sampling a novel word depended on the difference in relative frequency with which they had heard labels for their two options on past trials.

However, there remains a key confound in the current analysis of infants' sampling behavior across trials. Since the relative informativeness measure is computed in part based on infants' own sampling choices, this measure becomes highly correlated with infants' visual exposure to each of the novel objects across the sampling phase (a visual familiarity confound that was controlled across the training phase of the experiment). As a given object-label association becomes more informative across the sampling phase due to the fact that infants have not triggered a labeling event involving that object, infants' visual experience with the object also decreases relative to the other novel objects in the experiment. We thus designed a third experiment in which we aimed to address this confound by systematically manipulating informativeness of different object-label associations while controlling for visual experience.

Experiment 3

In Experiment 3, we sought to address the limitations of Experiments 1 and 2 by parametrically manipulating relative naming frequency while equating object exposure during the training phase. In order to manipulate relative frequency within a given participant, we adapted the design of Experiments 1 and 2 into four separate training and sampling blocks involving two novel object-label associations labeled at different frequencies. The relative frequency with which the two novel object-label associations were labeled was manipulated across blocks, such that some blocks contained a large discrepancy between the labeling information provided relative to the other (as in the training phase of Experiments 1 and 2, one object was labeled far more frequently than the other) and some blocks contained similar or equal labeling information for both novel objects during the training phase. The central question was whether the likelihood that infants selected a given object-label association depended on the relative frequency with which they had experienced the two object-label associations during the training phase. Analyses were registered prior to data inspection at the Open Science Framework (https://osf.io/khg7a).

Methods

Participants

Our target sample size was 50 19- to 21-month-old infants, based on a power analysis for our logistic mixed-effects analysis (described below) using simulations built with the simr package in R (Green & MacLeod, 2016). We conducted simulations for a range of possible values for our main beta coefficient of interest using random effects and residual variance estimated from a pilot sample of 10 participants, for both of the two main analyses described in the Results section. We found that a sample size of 50 participants would provide roughly 80% power to detect an effect of informativeness half of the size ($b = 0.4$) of that observed in two previous experiments ($b \sim 0.8$ in the informativeness analyses in Experiments 1 and 2).

Fifty-two 19- to 21-month-old infants (34 female) participated in the experiment (mean age =19.8 months, SD = .60, range: 18.9 - 21.1 months). An additional 15 infants were tested but excluded from the analyses due to experimenter error (n = 4) or infant fussiness (n = 11). We collected 2 additional infants beyond our target sample since families were contacted in batches larger than our target sample, to account for medium response rates and fuss-out rates typically observed in infant studies. All infants were English-learning, full-term infants with no vision or hearing problems and no exposure to a second language. Families were recruited from the Madison, WI community and received a gift for participating in the study.

Stimuli

In addition to the 4 novel objects and names from Experiments 1 & 2, we selected four unfamiliar objects from the Novel Object and Unusual Name Database (see Appendix A; Horst

& Hout, 2016) and four novel labels (*beppo*, *nillet*, *roozer*, *soma*). The two sets were randomly assigned to different training blocks and labeling frequencies across the experiment.

All auditory stimuli were spoken by a female native English speaker from Wisconsin in child-directed speech. The novel words were normalized for intensity and duration such that the duration of all novel words was equal. The words were presented in carrier phrases that were normalized for intensity.

Design & Procedure

The experiment was structured into four independent blocks. These blocks were designed to be shortened instantiations of the training and sampling phases from Experiments 1 and 2 involving only two novel objects and labels. Within each block, infants were exposed to two new object-label associations and subsequently given the opportunity to trigger additional learning events in a sampling phase. The key manipulation was that the relative frequency of one object-label association compared to the other was varied across the four blocks, such that children experienced object-label association frequencies of 4:0 (i.e. one object was labeled four times, while the other was never given a label), 4:1, 4:2, and 4:4 in one block each (Figure 4.8).

Sampling training. As in Experiment 2, infants completed four sampling trials at the outset of the experiment with familiar objects and labels (*ball*, *cat*, *cookie*, *shoe*), in order to familiarize infants with the gaze-contingent procedure.

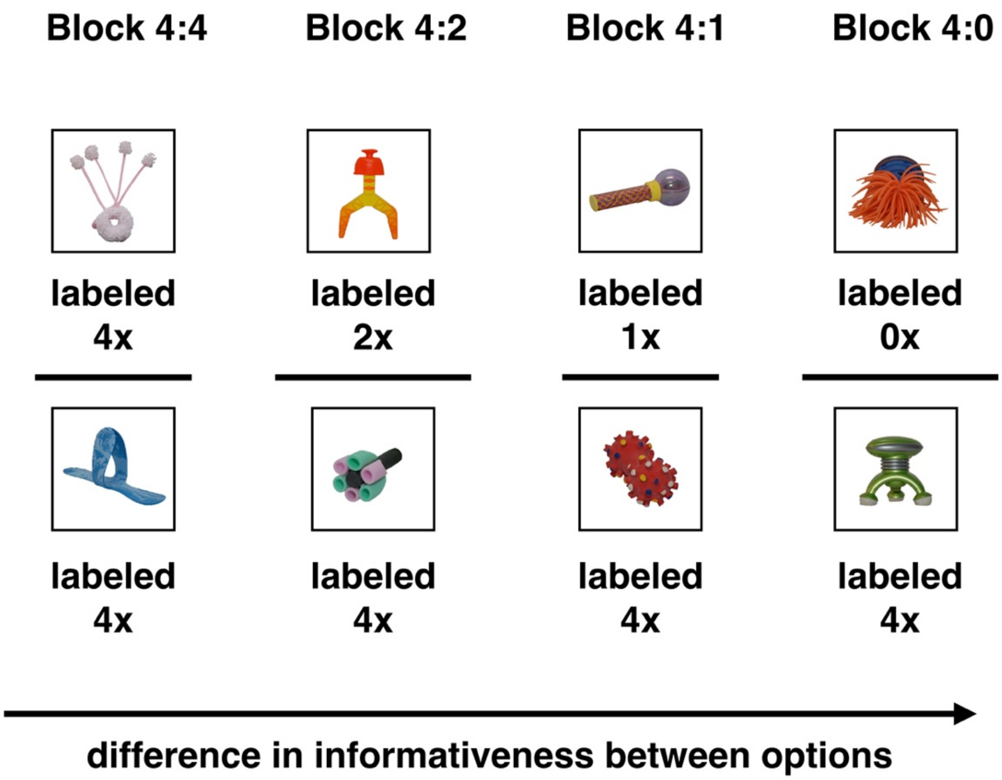

Figure 4.8. Manipulation of relative frequency of naming across training blocks. The order of blocks was randomized across participants.

Training Phase. Individual training trials were designed as in Experiments 1 & 2, such that on each trial, a single label was presented together with either a phrase labeling the object or a phrase directing infants' attention to the object without providing a label. Within a given block, each object occurred in isolation on four separate trials, for a total of 8 training trials per block (2 objects x 4 training trials). Additionally, each training block began with a familiar labeling trial presenting one of the 4 familiar items (*ball*, *cat*, *cookie*, *shoe*), selected at random, to reinforce throughout the experiment that the speaker was labeling the objects presented on the screen.

The key manipulation was the frequency with which each object occurred with its label over the course of a given block. Within each block, one object was always labeled on all four training trials, as in the high-frequency condition of Experiments 1 and 2. The number of labeling events for the other object was manipulated across blocks (Figure 4.8). In block 4:0, the second object was never labeled, and instead only appeared with attention-getting audio without label information on all four trials. In block 4:1, the second object was labeled once, as in the low-frequency condition of Experiments 1 and 2. In blocks 4:2 and 4:4, the second object was labeled twice and four times, respectively. Consequently, while all objects were seen equally often within a given block, the relative frequency with which the two objects were labeled varied across the four blocks of the experiment. The frequency role of the object pairs was counterbalanced across infants.

Sampling Phase. The design of an individual sampling trial was the same as in Experiment 2. At the end of the training phase of a given block, children saw two sampling trials involving the two novel objects from that block. The object locations were randomized across participants and blocks, and the location of the objects was reversed on the second trial compared to the first trial of a given sampling phase.

Word Learning Test. At the conclusion of the experiment, infants completed a 2-alternative forced-choice (2-AFC) task testing their knowledge of the novel words. We included this task as an exploratory measure of infants' learning, expecting some attrition/ fussiness given the age of our participants and the duration of the in-booth experiment.

The task was presented on a touchscreen monitor outside of the experimental booth, using a browser-based experiment created with jsPsych (de Leeuw, 2015). We introduced the task as a game in which participants were helping the experimenter find animals and toys. Infants were instructed to point towards the screen to make their choice, with the experimenter selecting the option that participants pointed towards. The task began with a short warm-up phase in which infants were asked to "catch" fish that popped up at various locations on the screen by pointing towards them. This phase familiarized infants with the procedure and helped encourage infants to interact with the events on the screen by pointing. Next, participants completed ten 2-AFC pointing trials, two involving familiar items (*cat*, *dog*) and eight trials testing infants' knowledge of the novel object-label associations. On each trial, two objects (either the two familiar or two novel objects) appeared side-by-side on the screen. Next, an audio stimulus was presented asking infants to point to a target object (e.g., *Where's the beppo?*). The eight novel labels were tested in random order. The location of the target was counterbalanced across trials.

Results

Sampling Preferences

We tested whether infants' likelihood of selecting novel-object label associations depended on the relative frequency with which two novel object-label associations were experienced during a given training block. Specifically, we asked whether infants' likelihood of choosing the more informative object-label association increased as the difference in relative frequency increased across the four training conditions. We conducted two main analyses, coding the condition manipulation as a continuous predictor in two different ways.

In the first analysis, we coded condition as a continuous predictor, based on the difference in labeling frequency between the two novel object-label associations in a given condition (i.e., condition 4:0 was coded as 4 – 0 = 4, block 4:1 was coded as 4 – 1 = 3, and so on). We fit a logistic mixed-effects model predicting the likelihood of infants' choosing the more informative object-label association object from condition (coded as described above). We included by-participant and by-item random intercepts and a by-participant random slope for condition. Contrary to our prediction, there was no significant relationship between the difference in labeling events during training and infants' preference for the more informative (i.e., less frequently labeled) object-label associations, $b = 0.01$, 95% Wald CI = [-0.13, 0.16], $z = 0.19$, $p = .85$ (Figure 4.9A). Similar results were obtained when the analyses were restricted to infants' first sampling opportunity in each block/ condition (Figure 4.9B).

In the second analysis, we coded condition analogously to the informativeness analysis in Experiments 1 and 2, in terms of the difference in informativeness of labeling events for the two novel object-label associations experienced during training. For example, condition 4:2 was coded as $\log_2(4/8) - \log_2(2/8) = 1$. Since $\log_2(0)$ is not defined, we set the relative frequency of the object-label association with a frequency of 0 to be half of the lowest possible relative frequency assuming a single labeling event (i.e., 1/16). Outcomes are similar for other possible choices for block 4:0. We fit a logistic mixed-effects model analogously to the first analysis, predicting the likelihood of infants' choosing the more informative object-label association object from condition (coded in terms of the difference in informativeness), including the same random effects structure as above. As in the first analysis, we found no relationship between the informativeness difference of infants' options and their preference for the less frequently labeled object, $b = 0.01$, 95% Wald CI = [-0.26, 0.28], $z = 0.06$, $p = .95$ (Figure 4.9C). As above, similar results were obtained when the analyses were restricted to infants' first sampling opportunity in each block (Figure 4.9D).

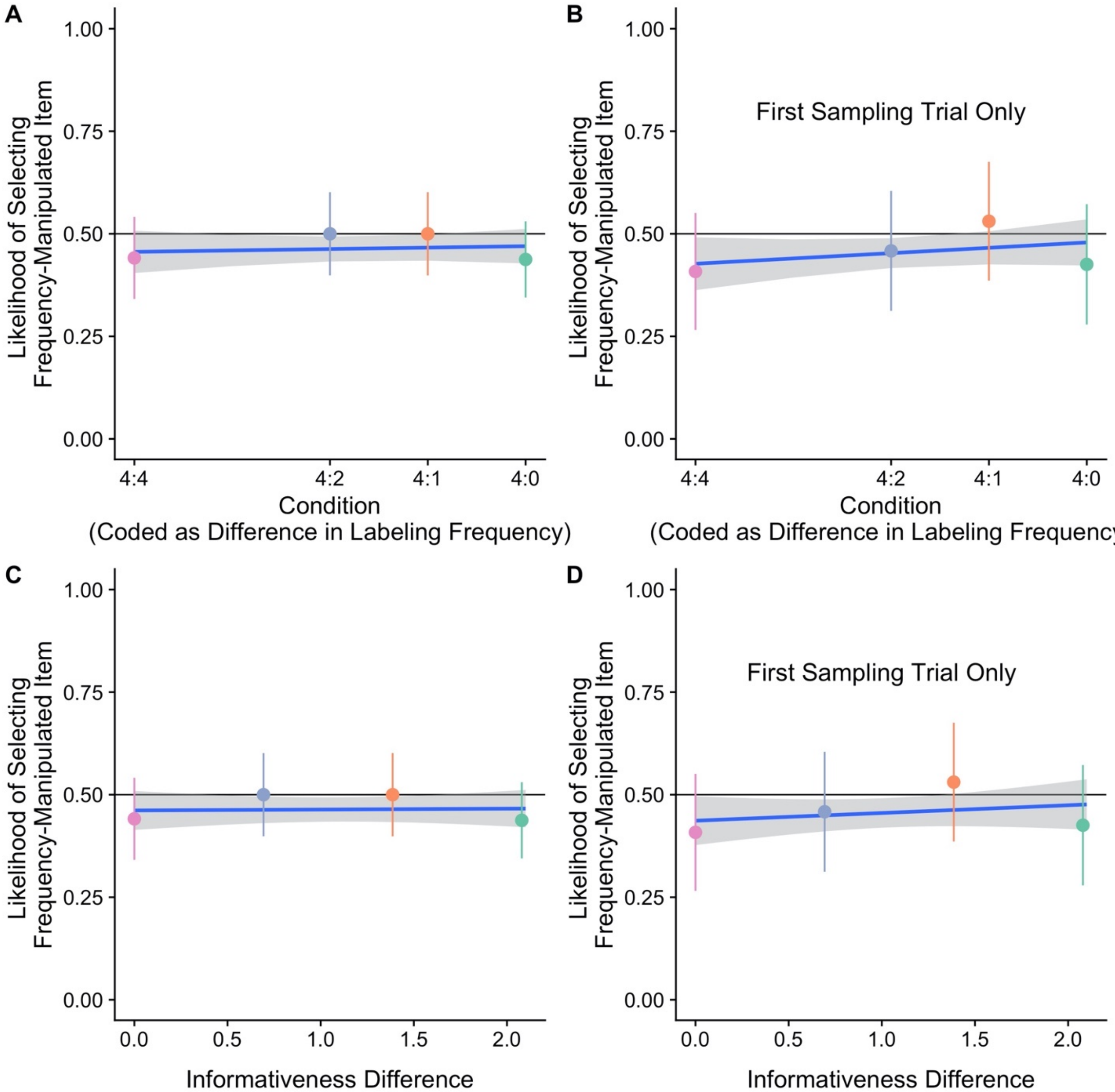

Figure 4.9. Overview over sampling preferences in Experiment 3. (A) Relationship between training condition (coded as the difference between frequency for labeling events for the two novel object-label associations) and preference for the more informative option (i.e., the item for which frequency of labeling was manipulated). (B) Relationship between training condition and preference for the more informative option, as in (A), based on infants' first selection in each condition only. (C) Relationship between training condition (coded in terms of the difference in informativeness between the two novel object-label associations) and preference for the more informative option. (D) Relationship between difference in informativeness and preference for the more informative option, as in (C), based on infants' first selection in each condition only. In each figure, error bars represent 95% CIs for each condition. The blue line represents the estimate from the logistic mixed-effects model, with error bands representing +1/-1SEs of the model fit.

Word Learning Test

Overall, infants were able to identify the correct object-label association above a chance level of 0.5, M = 57.7%, 95% CI = [51.3%, 64.1%], $t(43) = 2.41$, $p = .02$ (Figure 4.10A).

We first investigated whether children's accuracy was related to the total frequency with which they heard a given object-label association, combining their exposure across the Training Phase and the Sampling Phase. We fit a logistic mixed-effects model predicting infants' trial-by-trial accuracy from their frequency of hearing a given object-label association during the Training Phase, including by-participant and by-item random intercepts and a by-participant random slope for word frequency. Children were more accurate for items that were heard more frequently during the Training and Sampling Phase, $b = 0.14$, 95% Wald CI = [0.01, 0.27], $z = 2.15$, $p = .03$ (Figure 4.10B).

Next, we asked whether children's accuracy was specifically related to the frequency with which they had heard labels during the Training Phase and the Sampling Phase considered separately. We fit a logistic mixed-effects model predicting infants' trial-by-trial accuracy from the frequency of exposure during the Training Phase and the frequency of selection during the Sampling Phase, including by-participant and by-item random intercepts as well as by-participant random slopes for each of the frequency predictors. We found more frequent exposure during the Training Phase was marginally related to higher accuracy during the Training Phase, controlling for frequency of selection during the Sampling Phase, $b = 0.14$, 95% Wald CI = [-0.01, 0.29], $z = 1.89$, $p = .059$ (Figure 4.10C). The frequency with which infants selected object-label associations during the Sampling Phase did not significantly predict their test accuracy, $b = 0.12$, 95% Wald CI = [-0.23, 0.48], $z = 0.69$, $p = .49$ (Figure 4.10D).

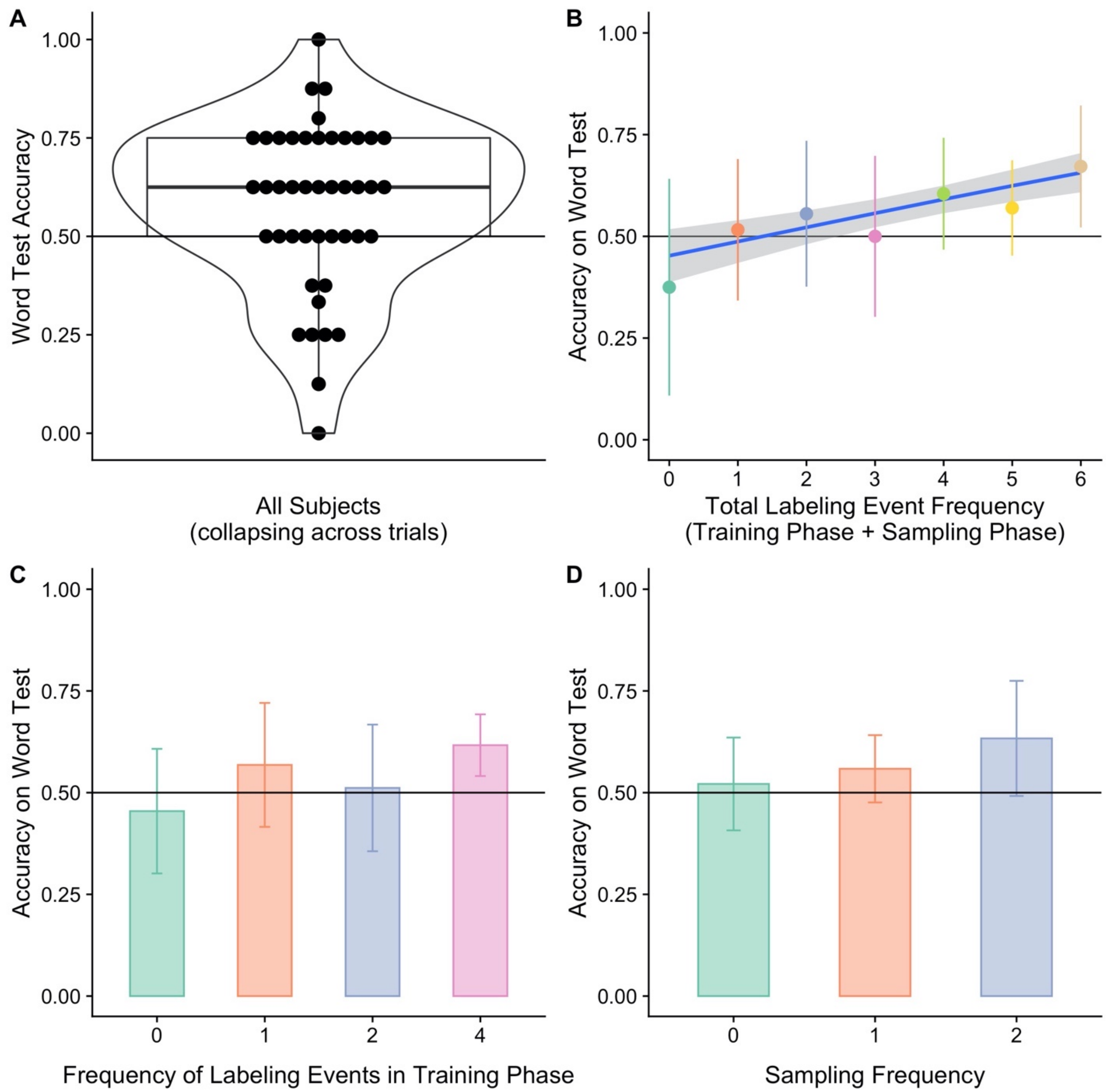

Figure 4.10. Overview of test accuracy in Experiment 3. (A) Overall distribution of participant accuracy, collapsing across all test trials. (B) Relationship between the total frequency of labeling events experienced prior to test (both during the Training and during the Sampling Phase) and word test accuracy. The blue line represents the (transformed) fit from the logistic mixed-effects model predicting participant accuracy from total labeling frequency; error bands represent +1/-1 SEs. (C) Accuracy at test depending on the frequency of labeling events for a given item during the Training Phase only. Error bars represent 95% confidence intervals. (D) Accuracy at test depending on the frequency with which a participant selected the item during the Sampling Phase. Error bars represent 95% confidence intervals.

Discussion

Contrary to our predictions, we found no evidence that the relative frequency of labeling events influenced infants' preference for sampling novel object-label associations in a gaze-contingent sampling task. Infants showed no overall preference for selecting novel object-label associations that were experienced less frequently during training, and the difference in frequency did not impact infants' preferences. Note that we also did not observe a significant preference for the less-frequently labeled object in any of the four blocks considered separately. In fact, the only condition in which infants showed a weak, non-significant tendency towards selecting the less frequently experienced object-label association (53.1%) was in the condition in which there was a 4:1 ratio between labeling events for the two object-label associations (but only when restricting the analysis to infants' first selections). This is the condition that matched the high- and low-frequency manipulation in Experiments 1 and 2, in which we observed a similar, small preference for sampling the novel object-label associations in the first block of sampling trials (~56% of trials). It is possible that there are "Goldilocks"-type effects in infants' sampling preferences, such that there is a "sweet spot" in which relative informativeness begins to inform infants' choices. However, given that our parametric manipulation of informativeness did not significantly modulate infants' choices, the present results suggest that relative frequency has little to no effect on infants' sampling in the current design.

One notable difference between the current design of the Sampling Phase and the Sampling Phase in Experiments 1 and 2 is that we restricted infants' choices to two sampling trials. This choice represented a compromise between the goal of limiting the potential of past sampling events to influence current sampling choices (as in the longer sampling phases in Experiments 1 and 2) and the goal of increasing power through repeated trials for each block.

However, one consequence of this design may have been that it encouraged infants to alternate between their selections, i.e. choose each object once. Indeed, infants showed a slight preference towards selecting both object-label associations in a given block (56.6% of blocks). This preference for alternating choices is not large, but may have been substantial enough to mask potential effects of informativeness. However, note that even if we restrict analyses to infants' selections on the first sampling trial of each block, we do not observe significant preferences in sampling, suggesting that infants' are not strongly drawn to the more informative option in this task.

Infants' demonstrated some knowledge of the novel object-label associations at the end of the experiment. While overall accuracy in a picture-pointing task was modest (~58% mean accuracy), infants were asked to track a relatively large number of novel object-label associations (8 total novel labels) across the task. Accuracy was generally predicted by the frequency with which infants' experienced a given object-label association, but not by the number of sampling choices for a given item alone. In other words, infants were not particularly more likely to learn objects that they selected during the Sampling Phase.

General Discussion

Across three experiments, we observed either weak or no evidence that infants are drawn to sample more informative object-label associations when given control over their learning input in a gaze-contingent design. In two initial studies, we found evidence that infants show a weak preference for selecting more informative object-label associations during their early sampling opportunities, and may be systematically influenced by the informativeness of their options across all of their sampling opportunities. However, in a well-powered third experiment with a within-subjects design that ruled out a potential confound in this analysis, we found no evidence that informativeness influenced infants' choices, and no evidence for preferential sampling in general. Thus, although infants' appear to be selecting object-label associations somewhat systematically in the extended sampling phases in Experiments 1 and 2, the results from Experiment 3 suggest their choices may not have been driven by the informativeness of the object-label associations, and that any effects of informativeness are likely to be small.

These experiments represent a first foray into using gaze-contingent eye-tracking to elicit sampling preferences from infants, and as such, our conclusions are limited by many of the design choices in our gaze-contingent method. Past research has successfully used explicit pointing (Begus et al., 2014; Begus & Southgate, 2012) and social referencing (Hembacher & Frank, 2017; Vaish et al., 2011) to elicit sampling choices from infants between one to two years of age. Our reasoning in developing a gaze-contingent eye-tracking method was that a more implicit, gaze-based method may reduce some of the response demands in more explicit tasks and allow for more sensitive measurement of infants' sampling preferences. Based on our current results, it is still an open question whether gaze-contingent approaches will prove fruitful in measuring infants' sampling preferences (though see e.g., Miyazaki, Takahashi, Rolf, Okada, & Omori, 2014; Tsuji, Jincho, Mazuka, & Cristia, 2020; Wang et al., 2012; Wass, Porayska-Pomsta, & Johnson, 2011 for related applications of gaze-contingency in infancy research). Two aspects of the current design in particular could be improved in subsequent projects. First, our design used a binary outcome, namely which of two object-label associations an infant triggered during the sampling phase. A design in which infants could collect information in a more continuous manner could provide more sensitive measurement of infants' preferences and provide more power in comparisons across condition manipulations. Second, in order to reduce the possibility that infants' choices would be driven by visual preferences alone, we obscured each object behind a screen prior to the gaze-contingent procedure. This likely increased memory demands for infants participating in the task, particularly as the number of objects appearing in different locations increased in Experiment 3. In chapter 5, we present a novel gaze-contingent sampling design that addresses both of these limitations by (a) allowing infants to continuously sample label information by fixating a given object and (b) reducing memory demands by displaying all object options on the screen throughout the sampling phase.

Another limitation of the current studies is that they present individual words in unambiguous word learning contexts. The manipulation of frequency across object-label associations may simply not have been strong enough to elicit a motivation to seek new information in infants (though it did successfully modulate older children's behavior in chapter 3). In particular, infants' early word learning environment is characterized by a high degree of uncertainty and ambiguity regarding word reference (Clerkin et al., 2017; Medina et al., 2011), and infants may be particularly motivated to seek new information following events where evidence is confounded or ambiguous (Schulz & Bonawitz, 2007; Vaish et al., 2011). Moreover, self-directed learning may be a particularly powerful precisely in circumstances when there is high uncertainty in the previously experienced input, so long as learners are able to successfully construct learning events that reduce uncertainty or aid in disambiguating past evidence (Coenen et al., 2019; Cook et al., 2011; Hidaka et al., 2017). Thus, in chapter 5, we test infants' information-seeking about new words in a context where there is higher uncertainty for some object-label associations than others, using a gaze-contingent method that measures infants' sampling preferences in a continuous fashion.

Chapter 5: Do infants sample to reduce uncertainty during word learning?

When infants encounter new words, they must inevitably be uncertain about the word's reference (Quine, 1960). Even if children can infer that a word refers to a concrete object in their environment, they face a complex many-to-one mapping problem. Having just encountered a novel word, which of the many objects in the environment does it refer to? Conversely, having just encountered a number of different words, which of those words refers to the object currently in focus?

Computational modeling work suggests that being able to control learning in a self-directed fashion can vastly simplify the problem of identifying novel words, to the extent that learners preferentially sample those words that are most useful to learning (Hidaka et al., 2017). When confronted with ambiguous or uncertain learning events, sampling object-label associations that have the potential to reduce the most uncertainty about reference has the potential to vastly improve learning speed and outcomes (Coenen et al., 2019; Settles, 2012). Previous research on category learning (Markant & Gureckis, 2014; Markant, Settles, et al., 2016), causal learning (Coenen, Rehder, & Gureckis, 2015) and word learning (Kachergis et al., 2013) provide support for the notion that adults make information-gathering decisions that aid learning by creating samples that often reduce uncertainty.

To what extent are infants and children drawn to events that aid in reducing uncertainty during learning? There is some evidence suggesting that even children are motivated to reduce uncertainty after experiencing ambiguous evidence (Cook et al., 2011; Legare, 2012; Schulz & Bonawitz, 2007) or surprising events (Stahl & Feigenson, 2015). For example, children spend more time exploring a novel box after experiencing ambiguous evidence regarding how to operate the box (Schulz & Bonawitz, 2007), and often create interventions on the box that isolate different potential causal factors in the operation of the box (Cook et al., 2011), i.e. that aid in reducing uncertainty about how the box works. A separate line of work has found that 11-month-old infants are more likely to explore novel objects after watching them occur in a surprising physical event, such as appearing to float in the air instead of dropping to the ground (Stahl & Feigenson, 2015). Infants' exploration in these situations often seems to be directed towards reducing uncertainty about the objects' properties, e.g. repeatedly dropping an object that just appeared to float in the air.

However, it remains unclear whether infants are motivated to reduce uncertainty during word learning. The experiments in chapters 2 and 3 provided evidence suggesting that children will systematically sample novel object-label associations, particularly words that will reduce uncertainty about new words (chapter 2). Are infants' also motivated to sample novel object-label associations that reduce uncertainty about object-label mappings? In the present experiments, we aimed to test infants' information-seeking strategies in an ambiguous word learning situation in which there was uncertainty with respect to the mapping between a given object and its label.

Experiment 1

Our goal was to design an experiment in which infants could selectively sample object-label associations that had different levels of uncertainty based on past evidence. In order to manipulate uncertainty, we presented infants with object-label information that provided consistent evidence for a specific label (i.e., each time infants saw the object, they heard the same label) and object-label information that provided inconsistent evidence for a specific label (i.e., each time infants saw the object, they heard a different label). The central hypothesis was that infants would seek out more information about previously inconsistently labeled objects, in order to resolve the uncertainty in their previous training exposure.

A second goal in the current design was to develop a gaze-contingent method that created more immediate gaze-contingency from the perspective of the learner and allowed for continuous sampling (rather than discrete sampling events, as in the studies described in chapter 4). In the current design, infants control the duration of their word learning events by fixating a given object, which triggers continuous labeling of the object until the infant shifts their fixation away from the object.

Method

Participants

Forty 17- to 19-month-old infants (25 female) participated in the experiment (mean age =18.0 months, SD = .65). An additional 11 infants were tested but excluded from the analyses due to experimenter error (n = 2), infant fussiness (n = 5), calibration or eye-tracking difficulties (n = 4). All infants were English-learning, full-term infants with no vision or hearing problems and no exposure to a second language. Families were recruited from the Madison, WI community and received a gift for participating in the study. We additionally collected information on children's receptive and productive vocabulary size using the MacArthur-Bates Communicative Development Inventory: Short-Form Level I (MCDI I Short Form) (Fenson et al., 2000).

Stimuli

The novel objects consisted of four unfamiliar objects from the Novel Object and Unusual Name Database (see Appendix A, Figure A2; Horst & Hout, 2016) and ten novel labels (*beevo*, *guffy*, *jeffa*, *kita*, *leemu*, *manu*, *noopy*, *roke, sibu*, *toma*). Each object was selected to be distinct in shape and color. The images of objects were presented against a white background.

All auditory stimuli were spoken by a female native English speaker from Wisconsin in child-directed speech. The novel words were normalized for intensity and duration such that the duration of all novel words was equal. The words were presented in carrier phrases that were normalized for intensity.

Apparatus and Stimuli Presentation

Participants were tested using a Tobii T60 XL eye tracker with a sampling rate of 60 Hz and displayed on a 20.5'' screen. The auditory and visual stimuli were presented in PsychoPy (Peirce, 2007). Communication with the eye-tracker and gaze-contingent stimulus presentation was controlled using the PyGaze toolbox (Dalmaijer et al., 2014). The infants sat on their caregivers' laps approximately 60 cm from the monitor in a sound-attenuated booth. To prevent caregivers from influencing infants' behavior, caregivers wore darkened glasses. A five-point calibration sequence was used to obtain a reliable track of participants' looking location (four corners plus center).

Design & Procedure.

The procedure consisted of three phases: a training phase, a sampling phase and a test phase.

Training Phase. During the training phase, infants heard novel labels paired with 4 novel objects. On each trial, an object appeared on the right or left corner of the screen in silence for 1000 ms. Then, the object rotated quickly to the right and left in a wiggling motion for 1200 ms. This wiggling motion was accompanied by speech labeling the object using one of two carrier phrases (e.g., *Look at the toma. That's a toma*. or *See the toma. It's a toma*.). Then, the object remained on the screen for an additional 1000 ms in silence before the trial ended and the next trial began after an inter-trial interval of 500 – 700 ms (jittered). The novel objects were presented one time each in random order on each block. Each block was followed by a short attention getter (i.e., a scene of a natural landscape accompanied by an attention-getting phrase, e.g., *Wow, look at that! That's cool!*). Infants saw a total of four training blocks, such that each object occurred four times in total.

The central within-subjects manipulation was the *consistency* with which each object was labeled across the training phase (see Figure 3.1). For two of the novel objects (*Consistent condition*), infants heard a consistent label across all four training trials. For the other two novel objects (*Inconsistent condition*), infants heard four different labels together with the novel object, i.e. infants heard a different label with the object on each training trial. This manipulation aimed to elicit a situation in which infants had more uncertainty about the object-label mapping for the two objects in the Inconsistent condition as compared to the two objects in the Consistent condition.

Four labels (*beevo*, *manu*, *noopy*, *toma*) served as the 'true' labels and were randomly assigned to one of the four objects for each participant. We then randomly grouped the objects into two sets of two object-label associations and assigned the resulting sets to the Consistent condition or the Inconsistent condition. The assignment of a given set of objects was counterbalanced across infants such that for each infant that saw two sets of object-label associations assigned to the Consistent condition and to the Inconsistent condition, another infant was assigned the same object-label association sets with the opposite condition assignment. The six remaining labels were randomly assigned to be heard as candidate words for the two objects in the Inconsistent condition. The order in which infants heard the labels across blocks in the Inconsistent condition was randomized across participants.

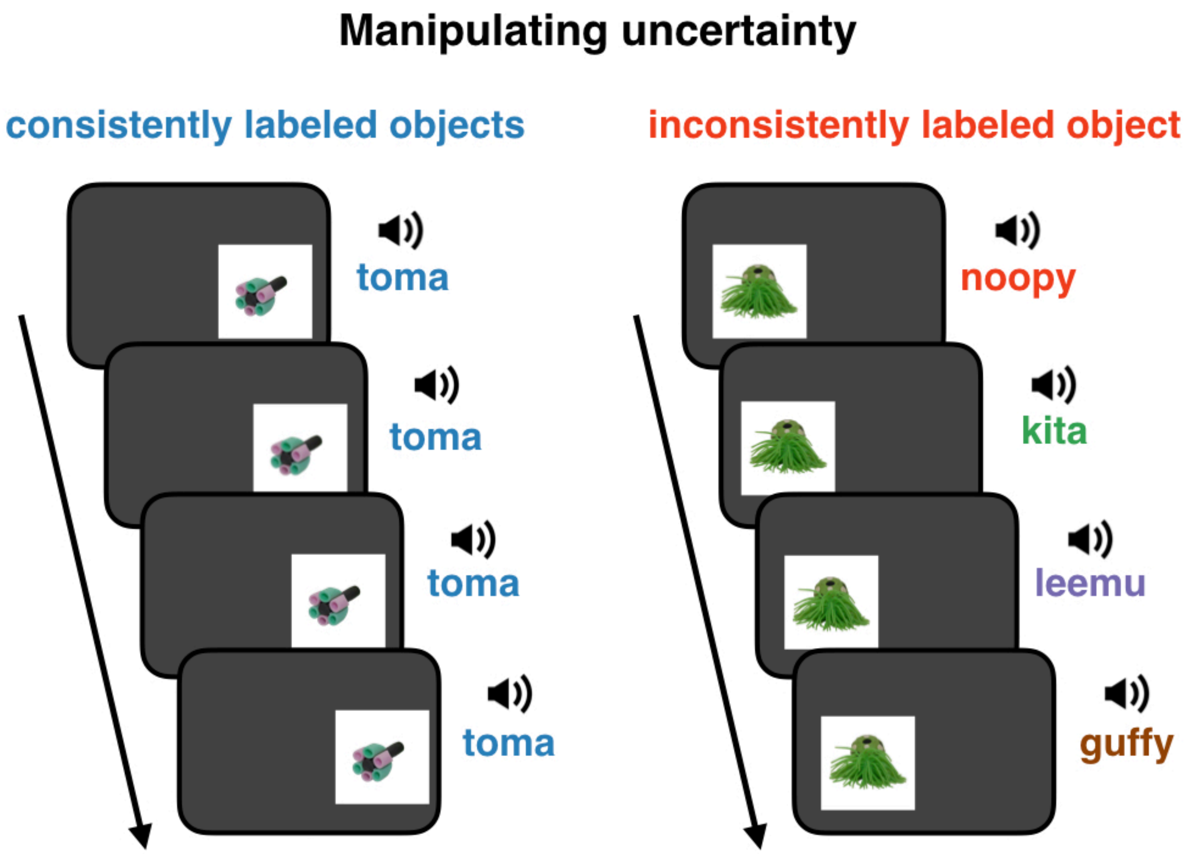

Figure 5.1. Consistent vs. Inconsistent condition manipulation during Training Phase.

Sampling Phase. In the Sampling Phase, infants had the opportunity to trigger labeling events involving the novel object-label associations in a gaze-contingent procedure. All objects were labeled consistently with a single label.

The Sampling Phase began with a trial with familiar objects to train participants on the gaze-contingent procedure. They then completed 4 sampling trials with the four novel object-label associations, with a 10-second attention getting trial (a picture of fish along with engaging music) occurring after the first two sampling trials.

On a given sampling trial, each of the four objects appeared one at a time in random order on the screen, with a short attention-getting "popping" sound accompanying the appearance of each object. The objects appeared at a pace of one object every 1000 ms. This introductory procedure was designed to draw infants' attention to each of the four objects and their locations on the screen. Next, the objects faded to grayscale and a pulsing green circle appeared at the center of the screen, to draw infants' attention to the center (central fixation; Figure 5.2). The pulsing green circle appeared for 1500 ms. Then, the gaze-contingent portion of the sampling trial began. Once infants fixated one of the objects for 300 ms, that object would appear in color and the accompanying label would begin to play in a continuous loop, with a 200 ms pause after the label ended. Each of these sampling events would continue for a minimum of 1000 ms (i.e., the duration of a full labeling event) until the infant looked away from the object (as registered by the eye-tracker). Once the infant looked away from an object and the sampling event ended, the infant could trigger another sampling event by fixating on any of the four objects for 300 ms (i.e., re-triggering the previously "sampled" object was allowed). Infants could continue to trigger sampling events until 20s had elapsed.

In order to make the sampling procedure more robust to eye-blinks and brief losses of data by the eye-tracker during infants' continuous fixations, we smoothed across the last 150 ms of looking (see e.g., Wass et al. 2013) by treating infants' current gaze location as the average looking location across the previous 150 ms of looking.

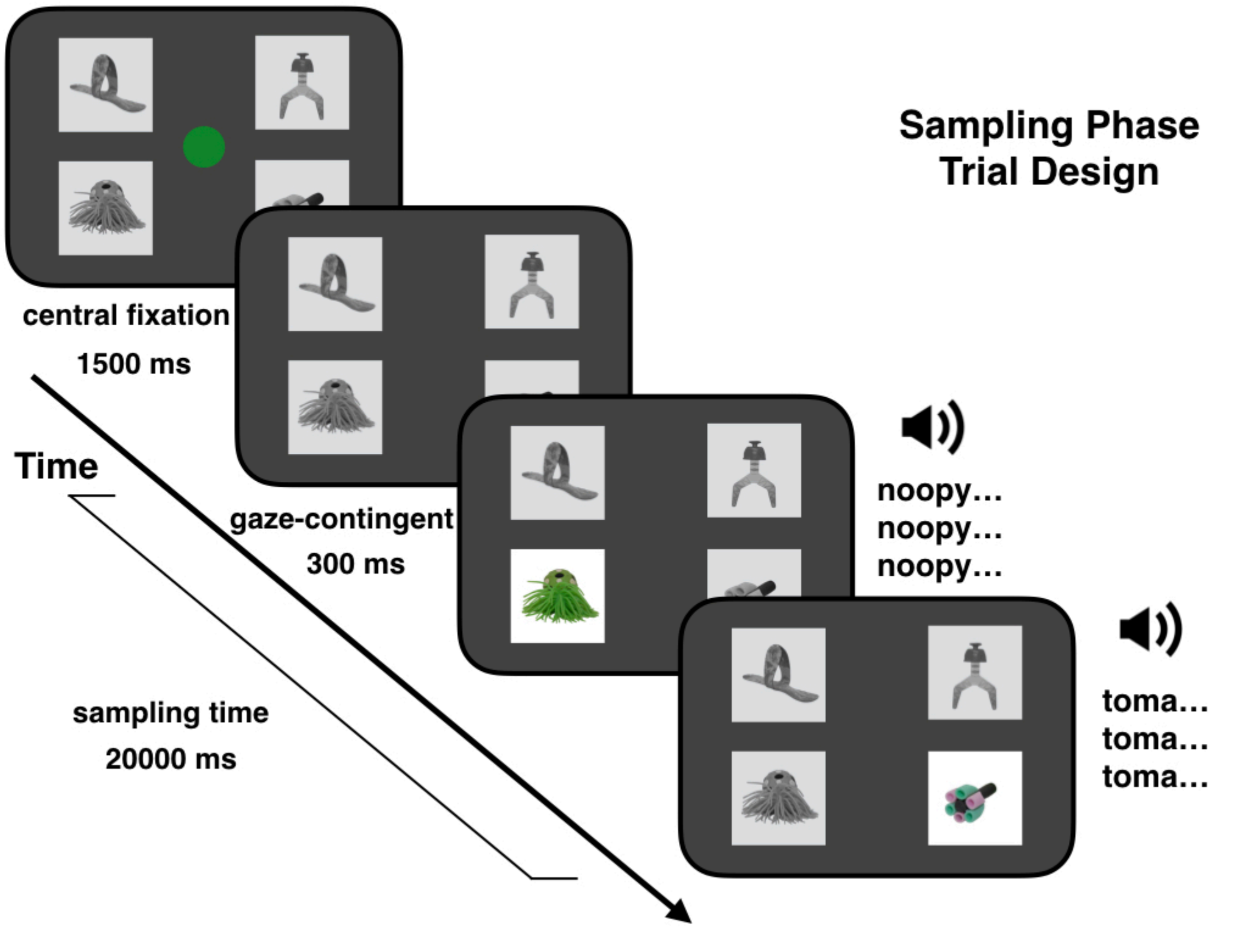

Figure 5.2. Sampling Phase trial design.

Results

Sampling Preferences

The central question of interest was whether infants would preferentially trigger sampling events for objects that had been inconsistently labeled during the Training phase. To test this question, we used the lme4 package version 1.1-21 in R (version 3.6.1) to fit a linear mixed-effects model predicting the average duration of time participants triggered an object from condition (Consistent vs. Inconsistent; centered). We included a by-participant random intercept. Degrees of freedom were estimated using the Kenward-Rogers approximation (Judd, Westfall, & Kenny, 2012). Infants sampled events involving objects from the Inconsistent condition for a longer duration than objects belonging to the Inconsistent condition, b = 2975, 95% Wald CI =

[903, 5047], $F(1,39) = 7.92$, $p = .008$ (Figure 5.3). There was a marginal interaction with sampling phase block, such that the effect became somewhat larger during the second half of the Sampling Phase (sampling trials 3 and 4) compared to the first half of the Sampling Phase (sampling trials 1 and 2), $b = 2140$, 95% Wald CI = [-335, 4616], $F(1,39) = 2.87$, $p = .098$.

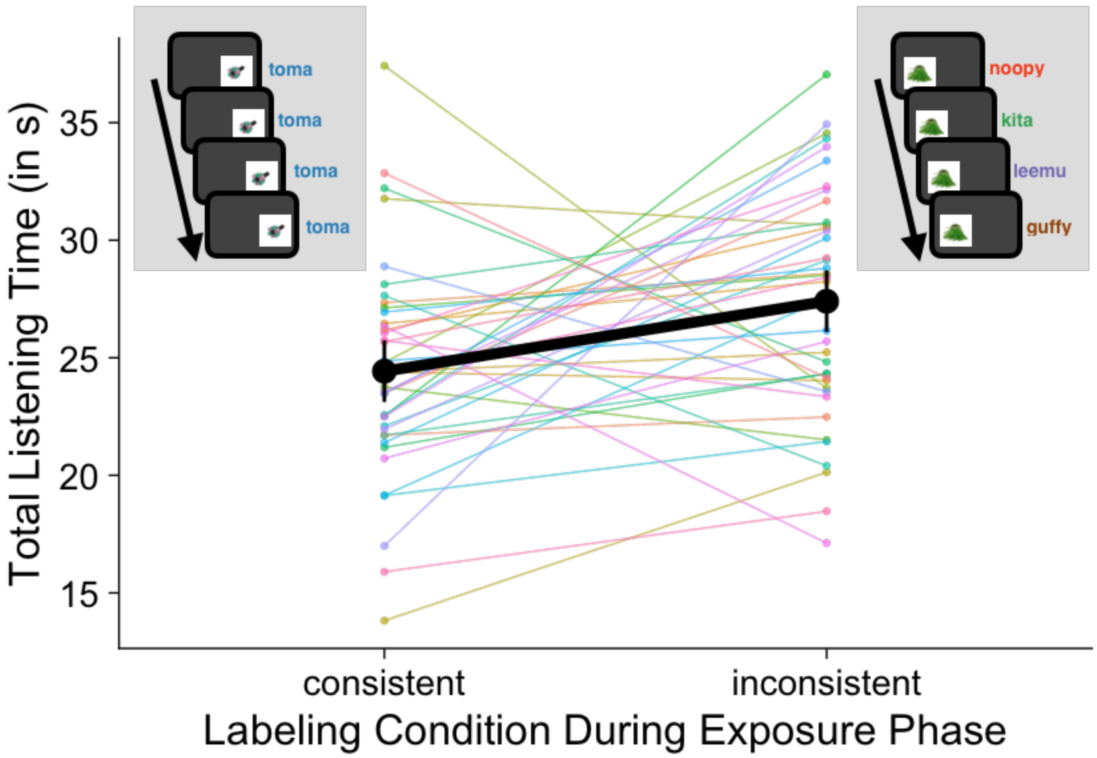

Figure 5.3. Overall listening time preference during the Sampling Phase in Experiment 1. Colored lines represent individual participants. Error bars represent 95% CIs.

We also considered alternative dependent variables in evaluating infants' preference for sampling object-label associations from the Inconsistent condition over object-label associations from the Consistent condition. We found a similar preference for object-label associations for the Inconsistent condition measured in terms of the total number of labels that infants triggered ($F(1,39) = 9.79$, $p = .003$; Figure 5.4) or measured in terms of the total number of distinct sampling events that infants initiated ($F(1,39) = 4.75$, $p = .035$).

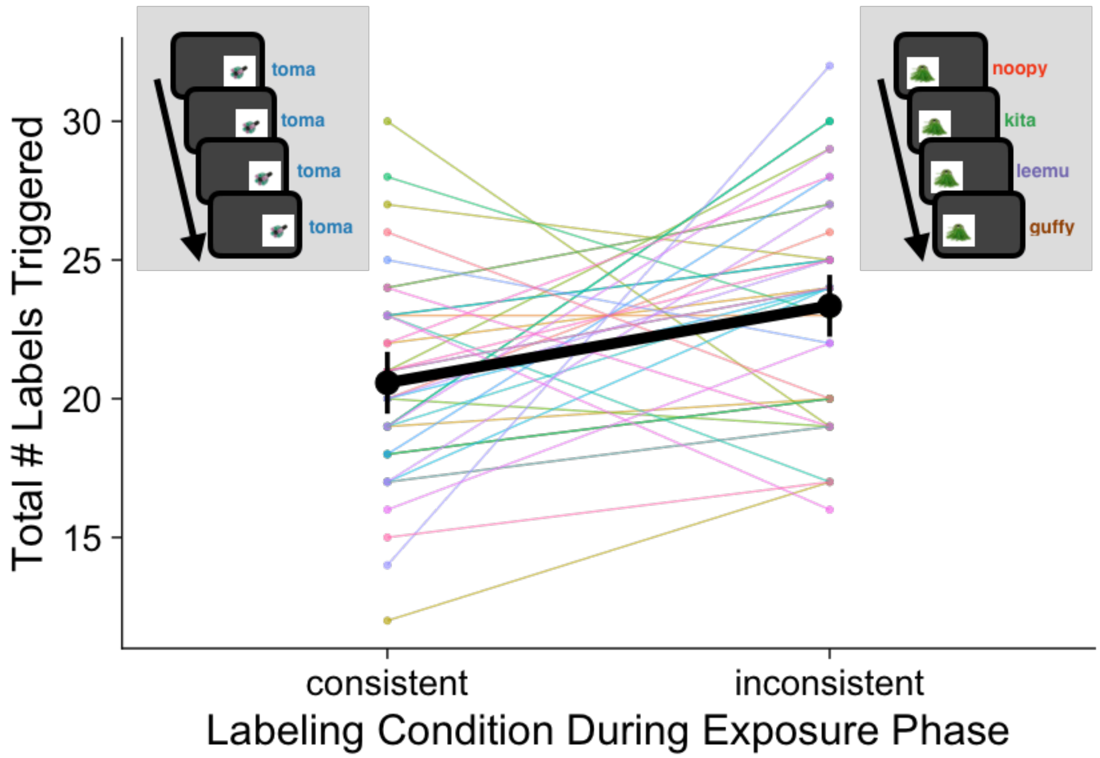

Figure 5.4. Total labels triggered during Sampling Phase in each condition in Experiment 1. Colored lines represent individual participants. Error bars represent 95% CIs.

Vocabulary Knowledge and Sampling Preference

In order to assess whether vocabulary knowledge was associated with different sampling preferences, we investigated the correlation between infants' preference for object-label associations from the Inconsistent condition over object-label associations from the Consistent condition (measured in terms of the difference in sampling duration for object-label associations belonging to the Inconsistent condition versus the Consistent condition) and infants' receptive and productive vocabulary as measured in the MCDI Level I (short form). We found that infants' receptive vocabulary was correlated with sampling duration preference, such that infants with the smallest receptive vocabulary showed the highest preference for sampling object-label associations from the Inconsistent condition, $r = -.46$, $p = .003$ (see Figure 5.5). We did not find a reliable correlation between sampling preference and productive vocabulary ($r = -.20$, $p = .22$).

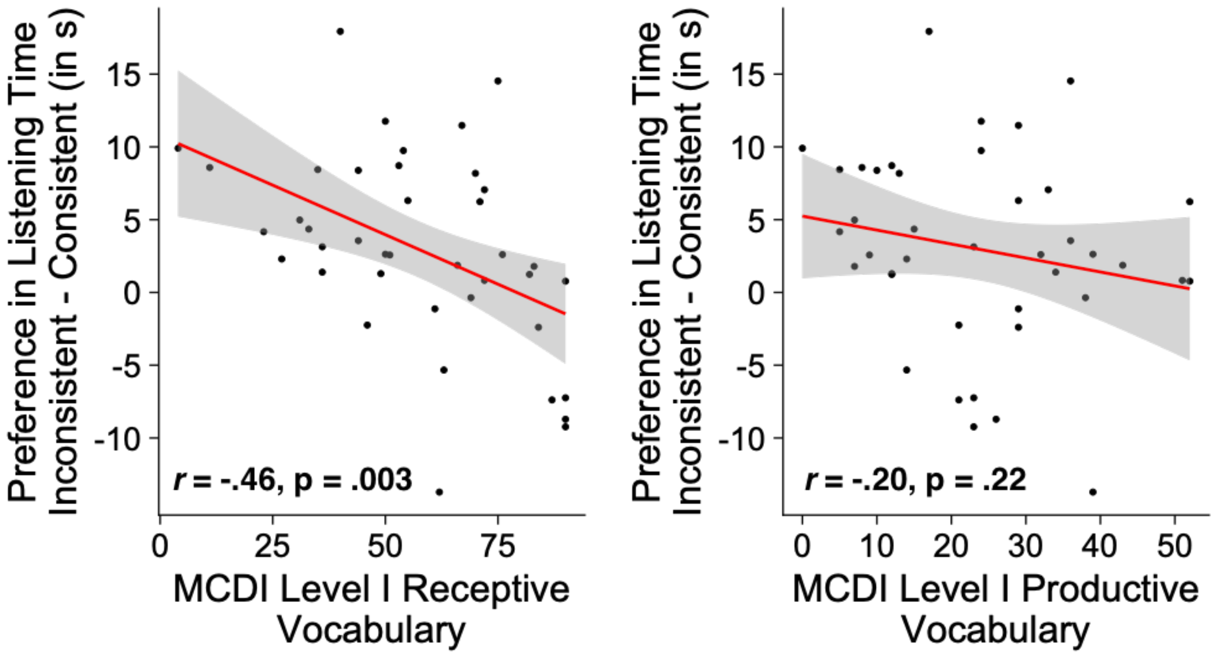

Figure 5.5. Relationship between listening time preference for the Inconsistent over the Consistent condition and infants' vocabulary size (MCDI Level I). Error bands represent +1/ -1 SEs.

Discussion

When given the opportunity to control their learning input, 18-month-olds preferentially sampled object-label associations for which they had previously experienced inconsistent labeling evidence compared to object-label associations for which they had heard consistent evidence. One possible interpretation of these results is that infants will seek to reduce uncertainty about novel object-label mappings – infants sampled words that helped resolve inconsistency in their past learning experience. These results are consistent with findings from other cognitive domains suggesting that infants are motivated to resolve uncertainty in previously experienced events (Schulz & Bonawitz, 2007; Stahl & Feigenson, 2015). Using this information-seeking strategy provides infants with a powerful means of resolving the ambiguity inherent in many word learning experiences (Medina et al., 2011; Quine, 1960).

In investigating individual differences in children's sampling preferences, we found that children with smaller receptive vocabulary sizes were more likely to prefer sampling previously inconsistently labeled object-label associations. This result was unexpected – *a priori*, one might have predicted the relationship to operate in the opposite direction, with children with larger vocabularies showing stronger information-seeking preferences. One speculative interpretation of these results is that children with smaller vocabularies require a larger amount of total evidence to overcome their noisy past experience. The 20s sampling trials provided infants with a substantial amount of new evidence about the words, potentially allowing rapid learners to quickly update their hypotheses about object-label associations. The children with smaller vocabularies, on the other hand, might be predisposed to update their hypotheses about object-label associations more slowly, explaining their stronger preference for object-label mappings with previously inconsistent evidence. This explanation – though speculative – suggests an intriguing direction for future research into how individual differences in vocabulary size might arise from differences in how children weigh past input.

Experiment 2

While Experiment 1 provides initial evidence that infants may prefer to sample object-label associations that help to resolve uncertainty, there is an alternative interpretation of the results in terms of auditory preference. The manipulation of labeling consistency in Experiment 1 is confounded with frequency of total label exposure in the training phase: participants heard the labels for consistently named objects in total more frequently than the labels for inconsistently named objects. Thus, infants' preference for object-label associations from the inconsistent condition during the sampling phase could be driven by the familiarity of the auditory stimuli alone, regardless of the relationship between the auditory label and its visual referent. In Experiment 2, our goal was to address this confound by manipulating labeling consistency while controlling for target label experience between the Consistent and the Inconsistent condition. The analytic approach was pre-registered in the AsPredicted format at the Open Science Framework (https://osf.io/uv5g8).

Method

Participants

Our pre-registered sample size was fifty-six 17-19-month-old typically developing infants. Based on a power analysis using the effect size found in Experiment 1 (d_z = 0.44), we estimated that our target sample size would give us 90% power to detect an effect of the same size. Fifty-eight 17- to 19-month-old infants (32 female) participated in the experiment (mean age = 17.7 months, SD = .56, range: 17.0 – 18.9 months). We oversampled our target sample size by two infants due to a slightly higher than expected response rate to our family recruitment efforts. Data from 24 additional infants was collected but excluded due to experimenter error (e.g., if parents were not properly blinded to the experiment with darkened sunglasses or other procedural errors; n = 3), poor or missing calibration (i.e., if fewer than at minimum 3 of the five calibration points are not satisfactorily calibrated; n = 1), parental interference (e.g., if the parent talked to the infant during auditory portions of the experiment; n = 1), or infant fussiness (e.g., if the experiment ended early because the infant was crying during the experiment or if the infant did not orient towards the screen during the Training or Sampling phase long enough to contribute useable data; n = 19).

Stimuli

We used the same visual stimuli as in Experiment 1. We used the same normalized audio recordings as in Experiment 1 with carrier phrases and labels recombined to follow the new labeling design in the Training Phase (see below).

Design & Procedure.

The procedure was identical to the procedure in Experiment 2, with one modification to the training phase.

Training Phase. During the training phase, infants heard novel labels paired with 4 novel objects. As in Experiment 2, on each trial, an object appeared on the right or left corner of the screen in silence for 1000 ms. Then, the object was accompanied by four 1500 ms phrases with a 300 ms pause between the onset of each phrase.

The key manipulation was whether the object was consistently labeled (Consistent condition) or inconsistently labeled (Inconsistent condition) on each trial (rather than *across* trials, as in Experiment 1). In the Consistent condition, the object was accompanied by one labeling phrase (e.g., *It's a manu* or *That's a toma*) along with three attention getting phrases (either *Hey look at that! – Do you see it? – That's cool!* or *Hey check that out! – Can you see it? – It's neat!*). In the Inconsistent condition, the listener heard 4 different labeling phrases (e.g., *Look at the roke* - *It's a manu – See the sibu – That's a jeffa*). The phrases used on a given trial in the Consistent and the Inconsistent condition were matched on the number of distinct linguistic types and total speech duration. This manipulation ensured that infants heard the "true" label assigned to each object on every trial, regardless of condition.

The novel objects were presented one time each in random order on each block. Infants saw a total of four training blocks, such that each object occurred four times in total. Crucially, given the new within-trial manipulation of naming consistency, infants heard each "true" label for each object (i.e., the label that reliably occurred when infants triggered a given object during the Sampling phase) an equal number of times in both the Consistent condition and in the Inconsistent condition during the Training phase (four times each). Labels were randomly assigned to objects as in Experiment 1 and the assignment of specific labels to a given object was counterbalanced across participants in the same manner.

Sampling Phase. The sampling phase was implemented identically to Experiment 1.

Picture Pointing Test. At the conclusion of the experiment, infants completed a 4-alternative forced-choice (4-AFC) task testing their knowledge of the novel words. The task was conducted using a touchscreen monitor outside of the experimental booth. Participants were introduced to the task as a game in which they were helping the experimenter find objects. Infants were instructed to point towards the screen to make their choice, with the experimenter selecting the option that participants pointed towards. The task began with a short warm-up phase in which infants were asked to "catch" fish that popped up at various locations on the screen by pointing towards them. This phase familiarized infants with the procedure and helped encourage infants to interact with the events on the screen through pointing. Next, participants completed six 4-AFC pointing trials, two involving the familiar objects from the experiment (objects: *cat*, *dog*, *duck, cookie*; target labels: *cat*, *dog*) and four trials testing infants' knowledge of the novel object-label associations. On each trial, four objects (either the four familiar or the four novel objects) appeared at one of four locations on the screen at random. Next, an audio stimulus was presented asking infants to point to a target object (e.g., *Where's the toma?*). The four novel labels were tested in random order. We included this task as an exploratory measure of infants' learning, expecting some attrition/ fussiness given the age of our participants and the duration of the in-booth experiment.

Results

Sampling Preferences

As in Experiment 1, our main hypothesis was that children would prefer to "sample" object-label associations belonging the inconsistent condition (i.e. items that were presented with multiple labels during the training phase) compared to object-label associations belonging to the consistent condition (i.e. items that occurred with a single, consistent label during the training phase).

To test this question, we implemented the same model as in Experiment 1. We fit a linear mixed-effects model predicting the average duration of time participants triggered objects from the fixed effect condition (Consistent vs. Inconsistent; centered), including a by-participant random intercept. Degrees of freedom were estimated using the Kenward-Rogers approximation (Judd et al., 2012). Infants had no significant preference for sampling events involving objects from the Inconsistent condition compared to objects belonging to the Consistent condition, b = -1161 ms, 95% Wald CI = [-3619 ms, 1298 ms], $F(1, 57) = 0.86$, p = .36 (Figure 5.6).

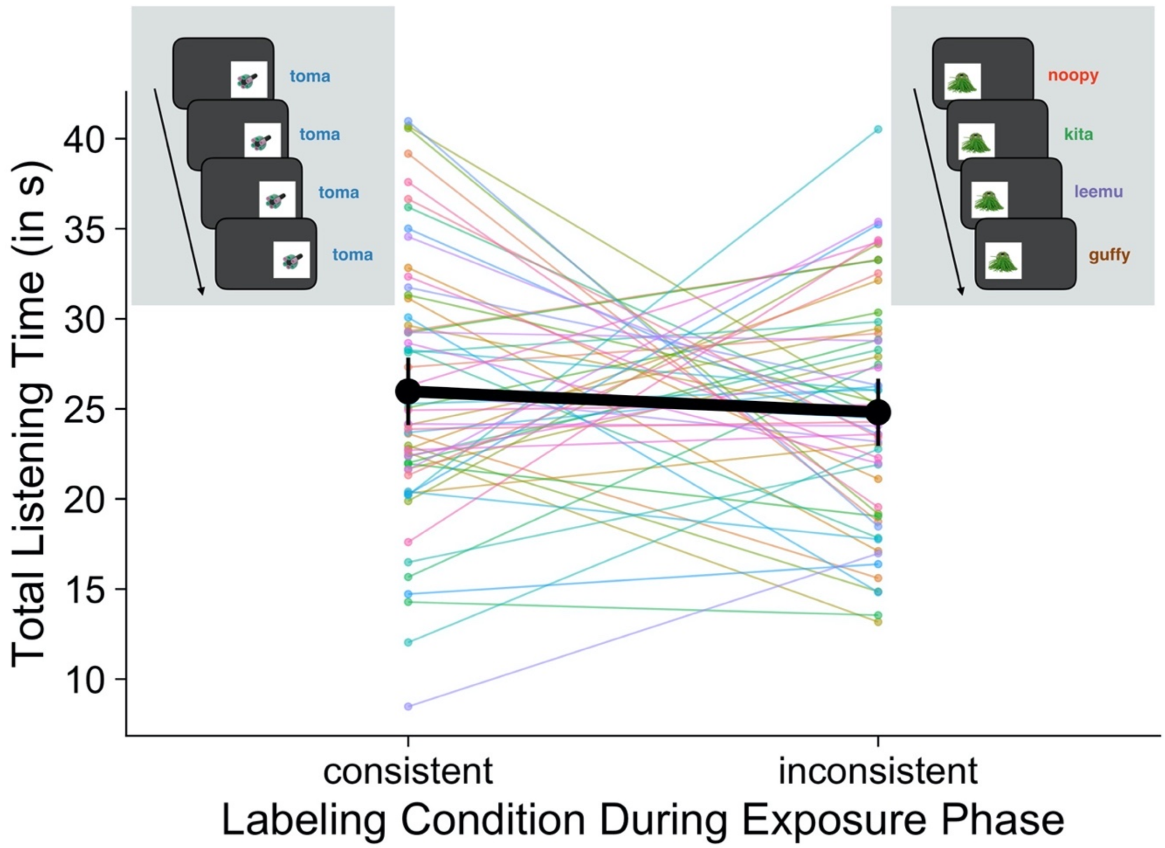

Figure 5.6. Overall listening time preference during the Sampling Phase in each condition in Experiment 2. Colored lines represent individual participants. Error bars represent 95% CIs.

We found similar results analyzing alternative dependent measures of sampling preference, namely the total number of distinct sampling events (b = -0.43, $F(1, 57)$ = 0.34, p = .56), and the total number of labels triggered for each condition (b = -1.03, $F(1, 57)$ = 1.03, p = .31; Figure 5.7).

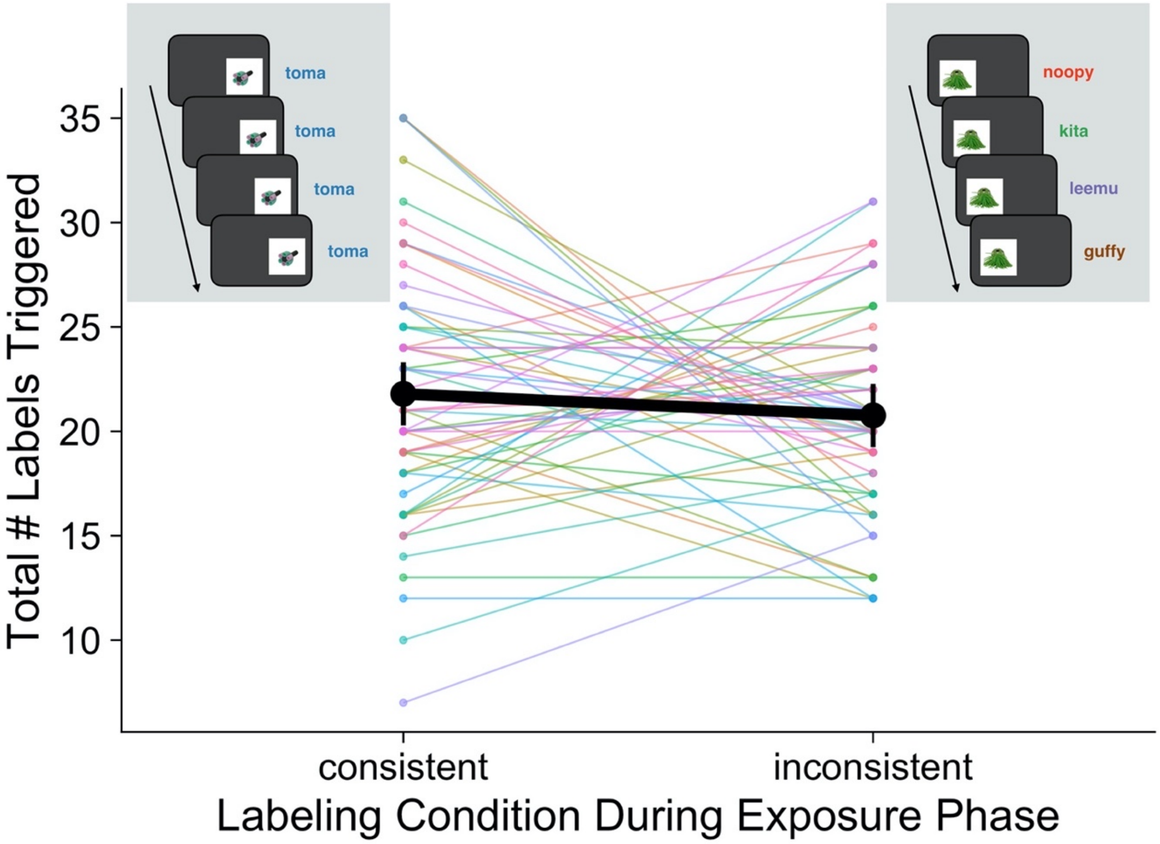

Figure 5.7. Total labels triggered during Sampling Phase in each condition in Experiment 2. Colored lines represent individual participants. Error bars represent 95% CIs.

Vocabulary Knowledge and Sampling Preference

As in Experiment 1, we investigated the correlation between infants' sampling preferences (the difference in sampling duration for object-label associations belonging to the Inconsistent condition versus the Consistent condition) and infants' receptive and productive vocabulary as measured in the MCDI Level I (short form). We found no significant relationship between infants' sampling preference and their receptive (r = -.09, p = .51) or productive (r = -.05, p = .69) vocabulary (Figure 5.8).

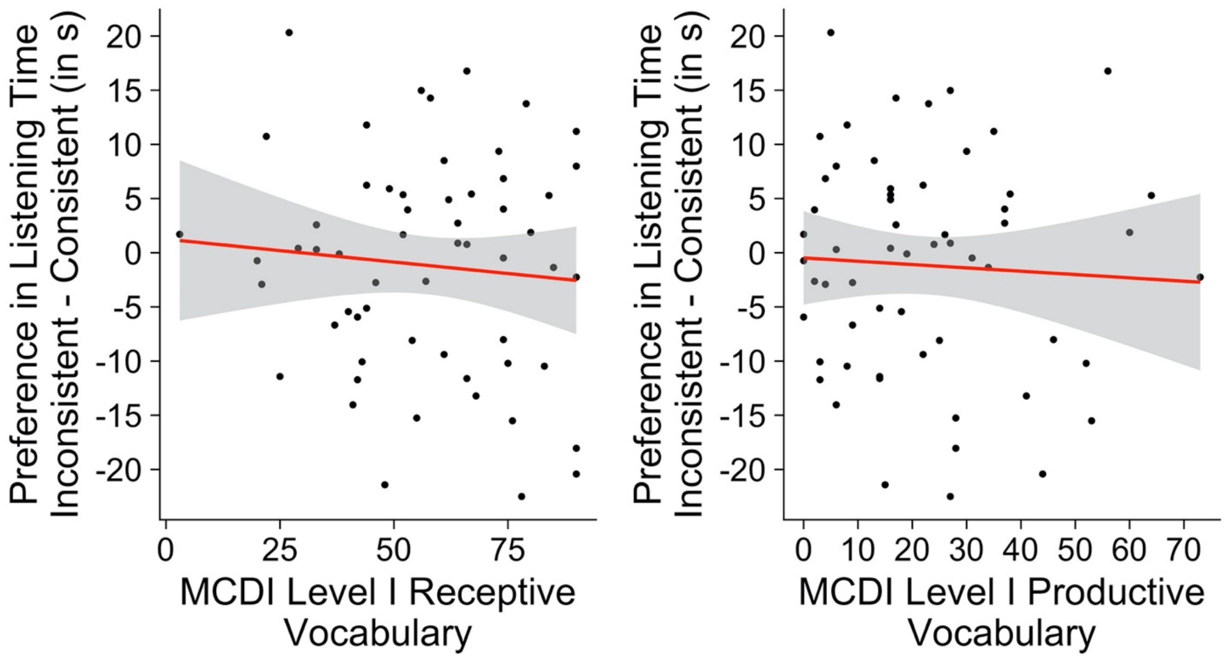

Figure 5.8. Relationship between listening time preference for the Inconsistent over the Consistent condition and infants' receptive and productive vocabulary size in Experiment 2. Error bands represent +1/ -1 SEs.

Picture Pointing Test

84.5% of participants (49 out of 58 infants) successfully completed the picture pointing task at the conclusion of the experiment. While infants typically successfully identified the familiar objects (M = 60.2%, 95% CI = [48.1%, 72.3%]; chance = 0.25), participants did not reliably select the novel target objects above chance (M = 29.9%, 95% CI = [22.4%, 37.3%]; $t(45) = 1.32$, $p = .19$).

Discussion

In Experiment 2, we sought to address a possible confound in Experiment 1 by controlling for auditory exposure across both the consistent and the inconsistent labeling condition during training. In a pre-registered experiment with sufficient power to detect an effect similar to that observed in Experiment 1, we found no preference for sampling object-label associations for which labeling information was inconsistent during training compared to object-label associations presented consistently during training. Unlike in Experiment 1, we also observed no consistent relationship between infants' sampling preferences and their receptive vocabulary size.

General Discussion

Are infants motivated to reduce uncertainty when learning novel words? The present studies sought to address this question by presenting 18-month-old infants with consistently or inconsistently labeled objects, and subsequently allowing them to control which objects they heard additional labeling information for. In Experiment 1, we found evidence consistent with the notion that infants are motivated to reduce uncertainty: 18-month-olds looked longer to objects that had been consistently labeled during training. Experiment 2 sought to control for an alternative explanation for the results in Experiment 1: since the "true" label for inconsistently labeled objects was heard less frequently (on only one out of four labeling events) compared to the "true" label for inconsistently labeled objects, perhaps longer looking to the inconsistently labeled objects was driven by an auditory preference. To address this possibility, we controlled for auditory exposure to the true label across consistently and inconsistently labeled objects in Experiment 2. Despite having sufficient power to detect an effect of similar size to Experiment 1, we found no significant preference for object-label associations that were inconsistently versus consistently presented during training in Experiment 2.

The present findings are consistent with three possible conclusions. First, the significant finding in Experiment 1 may be spurious or, relatedly, stem from a much smaller effect than was estimated from Experiment 1. The current experiments were well-powered for typical practices in infant research (Experiment 1: N = 40; Experiment 2: N = 58) and well-powered in general for detecting medium effect sizes in a within-participants design. However, it is nonetheless possible that the finding in Experiment 1 represents an overestimate of a much smaller true effect. Preliminary results from large-scale replication attempts investigating infant-directed preferences suggest that even putatively robust effects in infancy research may be smaller than has been previously thought (The ManyBabies Consortium, 2020).

A second possibility is that the effect estimated in Experiment 1 captured a true effect, but was (mainly) driven by differences in auditory exposure – infants in Experiment 1 were simply expressing a preference for auditory labels that had been heard less frequently during the previous training period. Once these differences in training exposure were accounted for in Experiment 2, there was no longer a significant effect. However, some patterns in the data are inconsistent with the conclusion that infant looking was driven solely by auditory exposure in Experiment 1. First, this explanation would predict that auditory preferences for inconsistently labeled objects (i.e., less frequently heard labels) in Experiment 1 should emerge early in the experiment, perhaps in the first or second sampling trial, and subsequently disappear once infants had begun to "even out" their auditory exposure across the labels. Contrary to this prediction, we observed that infants' preference for previously inconsistently labeled objects continued in the second half of the sampling phase of Experiment 1 (and was, if anything, stronger). Second, we observed a strong correlation between infants' receptive vocabulary size and their preferences for sampling object-label associations. While it is plausible that general auditory preferences may in some ways be linked to vocabulary size, it seems on its face more plausible that vocabulary size should be linked to individual differences in terms of how infants approach word learning – suggesting that infants' sampling preferences were at least to some extent connected to an attempt to learn the novel object-label associations.

A third possible conclusion is that infants' sampling in Experiment 1 was driven by a motivation to reduce uncertainty about the object-label associations, but that some other aspect of the design in Experiment 2 precluded participants from exhibiting this same tendency. We designed Experiment 2 such that, if infants did show a preference for object-label associations inconsistently labeled during training, this would constitute strong evidence for our hypothesis that infants were sampling words to resolve uncertainty. However, several aspects of the design may also have been more challenging for infants than the design in Experiment 1. In particular, the training time was substantially longer in Experiment 2 (approximately 3 minutes) compared to Experiment 1 (approximately 2 minutes). Moreover, it is possible that the mixed labeling design (in which multiple labels are presented for a single object on the same trial) employed in the current experiment may have been confusing for infants. While labeling the same object with different labels across several training trials, as in Experiment 1, has been successfully used to study infants' word learning under uncertainty (Vouloumanos & Werker, 2009), to our knowledge, designs in which many labels are presented for the same object on the same trial have not been extensively studied. It is thus possible that presenting multiple labels for the same object in close succession was simply overwhelming or distracting for 18-month-olds. It is notable that the fuss-out rate for participants in Experiment 2 (19 out of 82 infants, 23.2%) was more than twice as high as in Experiment 1 (5 out of 51 infants, 9.8%), possibly due to this combination of design decisions making the task in Experiment 2 more taxing for infants.

Further research is needed to disentangle these different possibilities. One path forward would be to consider simplifying the current manipulation of uncertainty, perhaps by reducing the number of different labels and objects, to rule out the possibility that the training design in Experiment 2 simply became too long and strenuous for infants. An alternative approach would be to consider manipulating the referential ambiguity of object-label associations (similar to the design in Chapter 2), since this has been shown to elicit social referencing behavior even in younger infants (Bazhydai et al., 2020; Hembacher & Frank, 2017; Vaish et al., 2011).

Many theories of child development posit that young infants are motivated to seek new information, particularly when it helps them reduce uncertainty about past and future events (Kidd & Hayden, 2015; Stahl & Feigenson, 2019; Twomey & Westermann, 2017). Motivations to actively seek new information about ambiguous information would be particularly powerful in the domain of word learning, given the ambiguity inherent in many word learning events and the vast lexicon that children must learn (Clerkin et al., 2017; Keijser et al., 2019; Medina et al., 2011). In both chapters 4 and 5, we found some preliminary evidence that infants may be motivated to create informative word learning situations, but the overall evidence in the present work is mixed. Future research will aim to disentangle the different possible conclusions left open in the current work, in order to develop a better understanding of how infants' curiosity supports their early word learning.

Chapter 6: Conclusions and open questions

1. Summary and discussion of the main findings

1.1. Children and adults sample words that reduce uncertainty.

Children are active, curious word learners. The results from chapters 2 and 3 support the conclusion that children and adults will sample novel words that reduce ambiguity about object-label associations (chapter 2) and that children will sample words that are more informative based on past training (chapter 3). The tendency to select informative options, observed in the study in chapter 3, emerged early, among children as young as 3 years of age. Moreover, in chapter 3, we observed evidence that children display graded sensitivity to the informativeness of past word learning events, with their preference for selecting informative object-label associations increasing as the difference in exposure frequency between their options increased. However, children's tendency to select novel words that reduced ambiguity emerged only at later ages in our sample: children began reliably selecting objects presented in ambiguous contexts at around 5 years of age in Experiment 2B in chapter 2.

This difference in the developmental trajectory of the sampling preferences observed in chapters 2 and 3 may stem from the increased complexity of learning words in ambiguous situations compared to learning words in unambiguous situations (but with varying frequency). In particular, children in the cross-situational word learning tasks in chapter 2 appeared to have difficulty learning words even when they were disambiguated across the learning trials – what items are the most "informative" for children after cross-situational training may vary substantially across age, as children's ability to track and remember information across trials increases. An important lesson from these findings is that what a learner may be curious to learn about will crucially depend on what they have learned in the past.

The discrepancy in the developmental trajectory observed in chapters 2 and 3 may also be related to different factors influencing children's sampling decisions. In particular, in the design of chapter 3, more informative sampling options also tended to be more novel overall, since informativeness of sampling options was manipulated in terms of frequency of exposure. In the experiments in chapter 2, we attempted to control for factors such as frequency of experience more closely, in order to ensure that any difference in sampling preference was due to a motivation to reduce ambiguity about novel words. It is possible that making systematic sampling choices in these situations, when there are fewer cues aligned towards the same choice, requires more sophisticated cognitive machinery, at least when making explicit choices (see section 2 for a review of relevant cognitive abilities undergoing development around these ages).

1.2. Sampling choices are related to word learning.

Across the studies in chapters 2 and 3, and to some extent in chapter 4, we also found evidence that sampling choices were related to later learning outcomes. Some of this evidence was indirect: in chapter 2 (2A), chapter 3, and chapter 4 (Exp 1), we found correlations between adults' and children's sampling choices and their word learning. Participants tended to perform better on words that were selected during the sampling phase. In the study in chapter 3, we also found more direct evidence that sampling was related to learning: children who were able to actively select their sampling options performed better than participants who experienced the same words during the sampling phase, but who had had different initial training on the new words compared to their yoked counterparts. This finding suggests that children's active selections were tuned to their past learning – sampling choices led to successful word learning in the Active condition, but were less optimal for learners with differing past experience. However, a consistent finding throughout the experiments was that making active selections, *per se*, was not necessarily more useful than passively experiencing those same choices. For example, in chapter 3, participants' performed better on words experienced during the sampling phase, regardless of whether they had selected those items (in the Active condition) or not (in the two passive conditions). Overall, the findings in chapter 2 and 3 suggest that children's sampling decisions are tuned to their past experience, and that children derive learning benefits from these sampling decisions.

In the design of the study in chapter 3, children's word exposure during the sampling phase was held constant, while children's initial training experience was manipulated. A useful direction for future research will be to investigate children's learning while varying their sampling experience. One interesting approach would be to compare children's learning for actively selected compared to randomly generated sampling sequences. If children perform better when trained on actively selected word input compared to randomly composed training input, this would further strengthen our preliminary conclusion that children make selections tuned to their learning needs.

1.3. Mixed or inconclusive evidence that infants sample words systematically

In chapters 4 and 5, we developed novel gaze-contingent methods to investigate whether infants between 17 to 21 months of age systematically seek new information and are motivated to reduce uncertainty about novel words. Our current findings present an inconclusive picture. In chapter 4, we found initial evidence that infants may have a weak preference for selecting informative object-label associations, and be sensitive to the informativity of different sampling options in general. However, in a well-powered replication aimed to rule out potential confounds in earlier studies, we did not find any evidence that 19-21-month-olds are sensitive to the informativeness of novel object-label associations when making sampling decisions. The present evidence suggests that – at least in the gaze-contingent design used in chapter 4 – children are not (or, at best, only weakly) influenced by the frequency of their past exposure to different object-label associations.

In chapter 5, we tested whether infants were motivated to reduce uncertainty about inconsistent evidence in past word learning events. In Experiment 1, we found that 18-month-olds showed a preference for sampling words that were labeled inconsistently during training. However, in a pre-registered, well-powered second experiment controlling for auditory exposure, we found no significant sampling preference for objects that had occurred with multiple words. The evidence is therefore currently inconclusive, and more research is needed to determine whether infants' sampling preferences in the current paradigm are robust in general, and if they are robust, whether they are driven by motivations to reduce uncertainty about past learning experiences. Future work will aim to address these questions by simplifying the complexity of the task (reducing the number of objects and labels) and investigating infants' sampling preferences under different uncertainty conditions.

2. Open questions and future directions

Together, our current results suggest that children are motivated to seek information about new words and can contribute to constructing learning opportunities that support their learning process, and mixed or inconclusive evidence with respect to infants' active information-seeking tendencies. In the sections below, we briefly outline outstanding questions and potentially fruitful avenues for research into children's active word learning.

2.1. Manipulating and modeling uncertainty and information-seeking

Different manipulations of uncertainty. The four lines of research presented in the previous chapters all shared a similar guiding question – how children make selections about what words to learn about next when confronted with options that vary in their potential informational value. Across experiments, the potential informativeness of different object-label associations was manipulated in three different ways (Figure 6.1): frequency (how frequently an object-label association was experienced; chapters 3 and 4); labeling consistency (whether an object was given the same vs. different labels across training; chapter 5); and ambiguity (whether which particular label went with which object was ambiguous vs. disambiguated across training; chapter 2). These manipulations were chosen because each has been shown to impact children's word learning (e.g., Hembacher & Frank, 2017; Roy, Frank, Decamp, Miller, & Roy, 2015; Vouloumanos & Werker, 2009) and each manipulation captures different learning scenarios that may induce states of uncertainty and/ or curiosity in the word learner. However, each manipulation also touches on different aspects of the word learning problem children face. For example, in the frequency manipulation, the primary challenge children face is tracking the individual object-label associations heard during training, but each object is always unambiguously labeled. On the other hand, in the ambiguity manipulations, different objects and labels are brought into direct competition with one another and must be disentangled in order to successfully learn the novel words. While exploring the implications of these different ways of manipulating uncertainty and word knowledge lies outside the scope of the present work, in future research, we plan to systematically study the impact of different uncertainty manipulations by modeling the different training regimes within the same computational model (e.g., Kachergis, Yu, & Shiffrin, 2012).

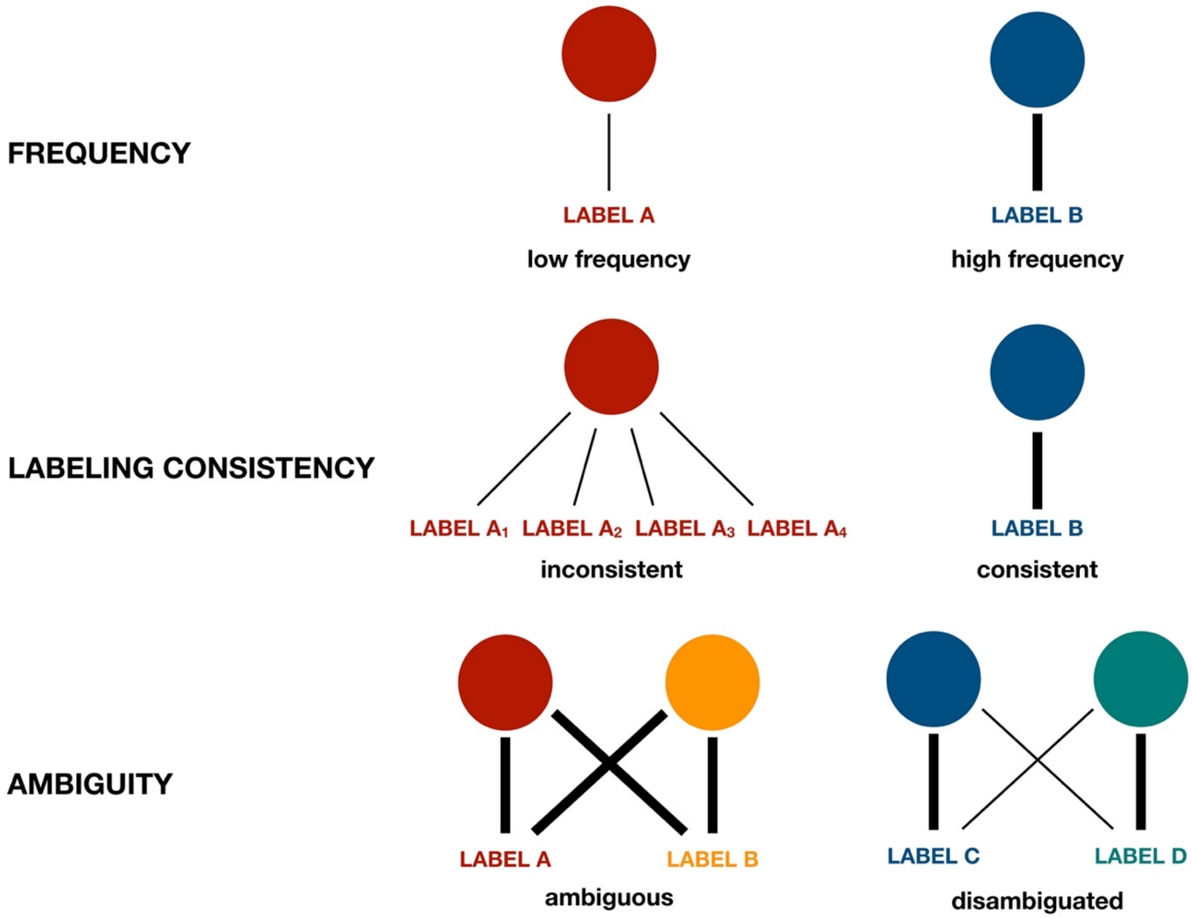

Figure 6.1. Different training manipulations of novel object-label associations used across experiments. Frequency of novel object-label pairs was manipulated in the experiments in Chapters 3 and 4. Labeling consistency was manipulated in the experiments in Chapter 5. Referential ambiguity was manipulated in the experiments in Chapter 2. The thickness of the lines connecting objects (colored circles) with labels represents the strength of the association between the object and label (i.e., the frequency of co-occurrence).

Connected to the goal of modeling different uncertainty states within one computational framework is modeling what types of decision rules learners implement when seeking new information about words. For example, one question is whether learners seek to reduce global uncertainty (i.e., focusing on simultaneously reducing uncertainty about all of the new object-label associations being learned) or local uncertainty (i.e., focusing on reducing ambiguity for a more limited set of items) in learning new words. Past work in category learning suggest that adults may favor reducing local uncertainty, focusing on comparisons between small sets of items when making their sampling decisions (Markant, Settles, et al., 2016). Our current data, in particular the sampling data collected in chapter 2, may represent an interesting test set for investigating whether local sampling strategies are also favored during word learning.

2.2. What developments in cognitive mechanisms support active learning?

A key question for future research is advancing our understanding of the cognitive mechanisms that allow children to develop flexible information-seeking skills. Two research areas that could be fruitfully integrated into current work on the development of active learning are research on the development of cognitive control and research on metacognitive abilities.

2.2.1. Cognitive control.

Information-seeking on most definitions is a goal-directed activity (Saylor & Ganea, 2018). In order for children to make goal-directed selections, this would seem to require (a) representing the underlying goal of the sampling action and (b) determining an action in the service of that goal. Children's developing ability both with respect to representing and maintaining a goal and acting prospectively in accordance with that goal should therefore have substantial influence on children's ability to flexibly and strategically collect information in the service of learning. Many of the cognitive skills associated with developing goal-oriented control over one's actions are studied under the concepts of executive functions and cognitive control (Munakata et al., 2011, 2012). This research emphasizes the importance of children's ability to represent abstract goal representations in the development of endogenous control over behavior (Munakata et al., 2011).

Two key transitions in children's cognitive control may be especially important to the development of information-seeking skills: the transition from reactive to proactive control - children's ability to plan in anticipation of expected outcomes – and the transition from externally-driven to self-directed control (Munakata et al., 2012). For example, one task that has been frequently used to document children's latter transition is the verbal fluency task. In this task, participants are instructed to name as many members of a category (e.g., foods) as possible in one minute. Children around the ages of four and five years will often name a limited number of category members and perseverate on already-named members, likely because they cannot yet effectively activate and switch between goal-relevant abstract representations such as subcategories (e.g., vegetables, fruits, etc.) (Snyder & Munakata, 2010). Similar skills in representing abstract information-seeking goals and inhibiting previously activated representations may be important in children's ability to strategically search for new information. Some initial findings support this prediction: Adams & Kachergis (2017) found that children tasked with learning a category boundary by actively sampling exemplars performed better on the category learning task when they had higher inhibitory control scores. Future research could elucidate the relationship between cognitive control and active learning by studying how transitions in children's executive functioning relate to changes in information sampling strategies (e.g., Ruggeri, Lombrozo, Griffiths, & Xu, 2016).

One interesting question that arises in the context of developing cognitive control is how well children can override the competing pull of particularly salient stimuli to make sampling choices that are most informative. One way to test this question in the current context would be to pit the salience against information gain in the design of a word learning task. In the experiments reported here, differences in saliency of competing choices were minimized in the design of the experiment and in the selection of stimuli. However, we still saw signs that children may sometimes make sampling choices that privilege more salient choices over more informative choices (in the sense of choices that provide more information about novel words).

For instance, in the practice phase of Experiments 2A and 2B in chapter 2, children often selected the familiar, (potentially) more salient known animal items over the (unknown) novel alien items. Future experiments could systematically investigate the development of children's preference for saliency over informativeness, and to what degree this tendency hinges on children's developing ability to exert endogenous control over attention.

2.2.2. Metacognition.

The development of metacognition – the ability to reason about one's own mental states – is another area that merits further integration with research on active learning. On some researchers' definition, metacognition may be a requisite skill for active learning. For instance, Saylor and Ganea (2018) define active learning as "involve[ing] the ability to identify gaps in one's knowledge" (p. 4). There are at least two lines of research that are relevant to understanding children's developing strategies for seeking new information: one focused on children's early explicit judgments about their own uncertainty and learning, and a second line of research focused on when and in what contexts infants show the earliest signs of metacognitive reasoning.

In the first line of research, past work has studied the development of explicit metacognitive judgments, in particular judgments of learning – explicit assessments of one's own learning and predictions about future performance – and confidence judgments – assessment of one's certainty about a response (Destan, Hembacher, Ghetti, & Roebers, 2014; Lyons & Ghetti, 2011). In general, children's estimates of future performance and confidence in their past learning exceeds later test performance (Lipko et al., 2009; Lipko, Dunlosky, Lipowski, & Merriman, 2012). However, there is also evidence that even preschool children are to some extent able to monitor their own learning, especially with increasing experience on a given task and when judgments are constrained (Lipko et al., 2012; Lipowski, Merriman, & Dunlosky, 2013). For example, after learning and being asked to recall (without feedback) a set of proper names for animals, five-year-olds were presented with a two-alternative forced-choice task in which they were asked which of two animals they would be more likely to recall later (Lipowski et al., 2013). Children were more likely to choose items that they had successfully recalled in the initial test compared to items they had failed to recall, showing sensitivity to their own learning. By the age of 6 years, children also devote more time to studying items which they think they have not learned well compared to items that they expect to remember when tested (Destan et al., 2014).

Children's judgments of their own confidence or certainty undergo significant development (Destan et al., 2014; Lyons & Ghetti, 2011). For example, in one experiment (Baer & Odic, 2019), children were asked to make approximate number judgments, deciding which of two groups of dots contained more items. Children's certainty about their own judgements was assessed in two different ways. In one task, children were presented with a choice between two potential trials, with varying ratios between the two dot groups in the two potential trials (e.g., one option was easier while the other required a more difficult judgment). In another, children were asked to make approximate number judgements for two trials that varied in difficulty (in terms of the ratio between dots), and then asked to decide which of the two they wanted the computer to "keep", i.e. the response that they had more confidence in. Children from 5 years of age tended to prefer both selecting and "keeping" easier trials with higher ratios between dot groups, while 3-4-year-olds showed no preference between easier and more difficult trials. Interestingly, children's performance on judgements of their own certainty tends to be highly correlated and undergo similar development across cognitive domains (Baer, Gill, & Odic, 2018). For instance, children's certainty assessments in approximate number tasks is correlated with their certainty judgements about facial cues to emotion (e.g., judging which of two facial expressions is "happier"). Thus, past work shows that explicit judgments about one's own confidence and learning undergo significant development across the preschool years, with children by the age of 5 showing a more consistent ability to evaluate their own certainty across domains.

A second line of research has investigated when infants show early signs of metacognitive reasoning using more implicit task designs. A commonly used approach capitalizes on infants' social referencing – infants' early-emerging tendency to look for cues from social partners to guide behavior (Feinman, 1982; Klinnert, Emde, Butterfield, & Campos, 1986; Walden & Ogan, 1988). Recent studies have demonstrated that infants and young children increase their degree of social referencing when confronted with more ambiguous or uncertain events (Bazhydai et al., 2020; Goupil et al., 2016; Hembacher & Frank, 2017; Vaish et al., 2011). For example, in a recent study, Goupil and colleagues (2016) demonstrated that 20-month-old infants begin to seek out the help of others in instances where they have a high degree of uncertainty. In this study, infants had to decide whether a hidden object was located in one of two locations. While one set of infants always made their decision with no outside prompting, another group of infants was given the option to either respond on their own on a given trial or to ask for help from their mother (after having been trained on how to do so in a previous task). Infants were roughly 3 times more likely to ask their mother for help on trials in which they were given no evidence about the object's location than when they were given the opportunity to observe the hiding event. Furthermore, on trials in which infants observed the object being hidden, their likelihood of asking for help depended on the difficulty of the task. The longer the delay between the hiding event and the decision about the object location, the more likely infants were to request their mothers' support. This experiment not only suggests that infants might be developing some type of meta-level understanding of their own uncertainty at a young age, but also illustrates the information-seeking flexibility and power children are acquiring: by developing an understanding of their own state of uncertainty, children now have the possibility of actively seeking out help and support from others in the service of reducing their uncertainty.

2.3. What does active learning look like "in the wild"?

One intriguing question for future research is what infants' and children's sampling strategies look like in more naturalistic settings. Across all of the current experiments, children received immediate, contingent supervised training in response to the events they initiate – every time they triggered an event, they heard a label for a distinct item. This kind of contingent training affords children a great deal of direct control over their experience. How representative is this kind of direct control of children's language experience in general? Are there analogues to this kind of experimentally manipulated active control "in the wild"?

One key to answering this question is understanding the dynamics between children and their caregivers. There is a great deal of research indicating that parents adapt their behavior, in particular their linguistic utterances, in response to their children's view, attention, and knowledge (Gros-Louis, West, & King, 2014; Matatyaho & Gogate, 2008; Trueswell et al., 2016; Y. Yu et al., 2018). Early caregiver-infant interactions are often characterized by contingent responses from caregivers to infants' vocalizations that support infants' early vocal learning (Bornstein, Putnick, Cote, Haynes, & Suwalsky, 2015; Goldstein, King, & West, 2003; Goldstein & Schwade, 2008; Gros-Louis, West, Goldstein, & King, 2006; Gros-Louis et al., 2014). Both experimental work (e.g., Horst & Samuelson, 2008) and more naturalistic investigations of caregiver-infant interactions (Yu & Smith, 2012) during word learning suggest that infants learn new words best when objects are labeled contingently on children's attention being directed towards the labeled object (so-called "follow-in" labeling). Moreover, caregiver's responses to children's early vocalizations and pointing behaviors are predictive of later language development (Gros-Louis et al., 2014; Wu & Gros-Louis, 2014, 2015). In general, the contingent nature of a large proportion of children's early language experience provides children with ample opportunities to actively probe their environment and expect informative answers – often delivered by caregivers.

However, two substantial caveats are warranted, given the state of current research on caregiver-child interactions. First, much of the research cited above is collected from WEIRD (Henrich, Heine, & Norenzayan, 2010) communities and the nature of parent-child interactions can vary widely across different communities and social contexts (Casillas, Brown, & Levinson, 2019; Heath, 1983; Hoff, 2006). Second, the fact that there is substantial variability even *within* communities in caregivers' responsiveness to infants' attention and information-seeking behaviors (e.g., Gros-Louis et al., 2014; Rowe & Goldin-Meadow, 2009) raises the question of whether infants adapt their information-seeking behavior to different social environments. An intriguing direction for future research is whether – and how - the nature of children's active learning changes in response to the particularities of their social and linguistic environment.

2.4. What is the right mix of active and passive learning?

There has been a wave of excitement for the possible benefits of allowing learners to control their input (Castro et al., 2009; Gureckis & Markant, 2012; Markant, Ruggeri, et al., 2016; Saylor & Ganea, 2018; Settles, 2012). One might think that active learning is simply inherently better than forms of learning where the learner cannot directly control their input (though see our results in chapter 3). While many studies have focused on the benefits afforded by active learning, most also find limits on the contexts in which active learning leads to superior learning outcomes (MacDonald & Frank, 2016; Markant & Gureckis, 2014). When is active learning most effective, and what is the right mix of active and passive learning?

2.4.1. Sampling dilemmas.

One useful way to conceptualize this question is to think about the composition of different samples that might result from active selection of the input versus passive exposure to input. Every sampling strategy has benefits and drawbacks, a problem Klaus Fiedler has termed the "ultimate sampling dilemma" (Fiedler, 2008). The key insight is that the usefulness of a given set of learning experiences will depend on the learning goal currently in focus – the sample that is optimal for one learning goal will not necessarily be the most optimal for other goals. To illustrate the inherent tradeoffs in any given information-seeking strategy, consider the following example. Imagine that a child is learning about the category *dog* and the category *cat*. To do so, they get to pick out different animals in their environment and ask, "Is this a cat or a dog?" How should they go about collecting information about these categories? The answer to this question will crucially depend on the problem the learner is trying to solve.

One potential goal is to learn to optimally *discriminate* dogs from cats, i.e. to learn the precise boundary between the two categories. A strategy that learners could pursue with this goal in mind is to focus on the problem cases – the cases where it is difficult to distinguish cats from dogs. This strategy leads learners to select exemplars that have the highest uncertainty, given their current priors about the two categories. In this example, learners might choose to gather information about dogs that look similar to cats and cats that look similar to dogs. While this strategy may help a learner figure out where exactly the line is between a cat and a dog, it also means the learner is exposed to a somewhat odd, unrepresentative set of cats and dogs. If these are the main examples of dogs and cats that a learner experiences, and they attempt to induce what the average dog or the average cat looks like from these exemplars, this will lead to representations of the categories dog and cat that do not match the natural statistics of their environment. For instance, dogs at the category boundary to cats are likely smaller than the average dog and likely have shorter snouts than the average dog.

On the other hand, if a learner's goal is to form general expectations about how dogs and cats (along with their respective properties) are distributed in the world, what seems optimal for the learner in this instance is to encounter exemplars of the categories according to their natural occurrence in the environment. It might be most useful, then, for the learner to absorb the distributional statistics of dogs and cats in their environment, without distorting those statistics by intervening on their learning environment. In other words, learners might benefit most from passive learning experience.

2.4.2. Representing the hypothesis space and active sampling.

A related issue is that generating samples appropriate to one's learning goals often requires having a good representation of the underlying hypothesis space – to the degree to which there is a mismatch between learners' representations and the underlying structure to be learned, active learning may become a less useful strategy (Gureckis & Markant, 2012). For example, in Markant & Gureckis (2014), when learners' task was to learn a category boundary that did not match their expectations– a category rule that required integrating across two dimensions rather than a (typically more expected) one-dimensional category rule –, receiving randomly selected category exemplars proved to be as effective as being able to actively select exemplars. Relatedly, learners in active learning tasks often begin with suboptimal sampling strategies before honing in on optimal choices as their understanding of the task grows (Markant & Gureckis, 2014; Shanks, Tunney, & McCarthy, 2002).

2.4.3. Combining active and passive learning.

One way to circumvent some of the issues with "pure" active learning strategies outlined above is to incorporate periods of both passive exposure and active sampling into learning experiences. Passive exposure to the natural statistics of the environment may help learners gain an understanding of the underlying structures, while active learning allows them to intervene and test hypotheses generated based on periods of passive observation. For example, using the same category learning task as Markant & Gureckis (2014), MacDonald & Frank (2016) have found that first giving learners passive exposure to category exemplars followed by a period of active selection leads to better learning than when the order of these conditions is reversed (i.e., active learning first, passive learning second). A similar notion is implemented in machine learning algorithms that combine active learning techniques with supervised or semi-supervised training regimes (Settles, 2012). Note that the design of all of the experiments reported in chapters two through five follow the "passive learning first, active learning second" pattern: these experiments all begin with an initial (passive) training phase, followed by an active sampling phase that tests participants' information-seeking strategies. Future work could delve deeper into the ways in which active and passive learning complement each other by probing the effectiveness of different combinations of active selection and passive exposure periods for children's learning.

3. Conclusion: Children as active learners in an interactive world

In the spirit of the information-seeking strategies at the heart of the present work, it seems fitting to conclude not by framing active learning as an answer, but as a question that should motivate further inquiry into the dynamics of early learning environments. Re-conceptualizing infants and children as active learners has powerful consequences for how we think about development. It brings into focus children's own agency in the learning process, and the ways in which curiosity-driven mechanisms help children discover structure in their environment, not only in language development, but across cognitive domains. However, it is important to recognize that children's information-seeking is tuned to a social environment that supports and rewards children's curiosity. Children's information-seeking strategies only make sense in the context of interactive environments that are responsive to children's questions – through socially contingent responses, through caregivers adapting their responses to children's learning needs, through teachers structuring children's early learning environment in guided play, etc. A fundamental challenge for future research will be to understand the interconnected dynamics of children's active learning strategies and the adaptations in environmental structure that make those information-seeking strategies effective for learning. Understanding children as active learners should not make us mistake them for isolated learners (like the children imagined in our cave thought experiment from chapter 1) – active learning works best in a social environment that supports curiosity.

Appendix A: Visual stimuli

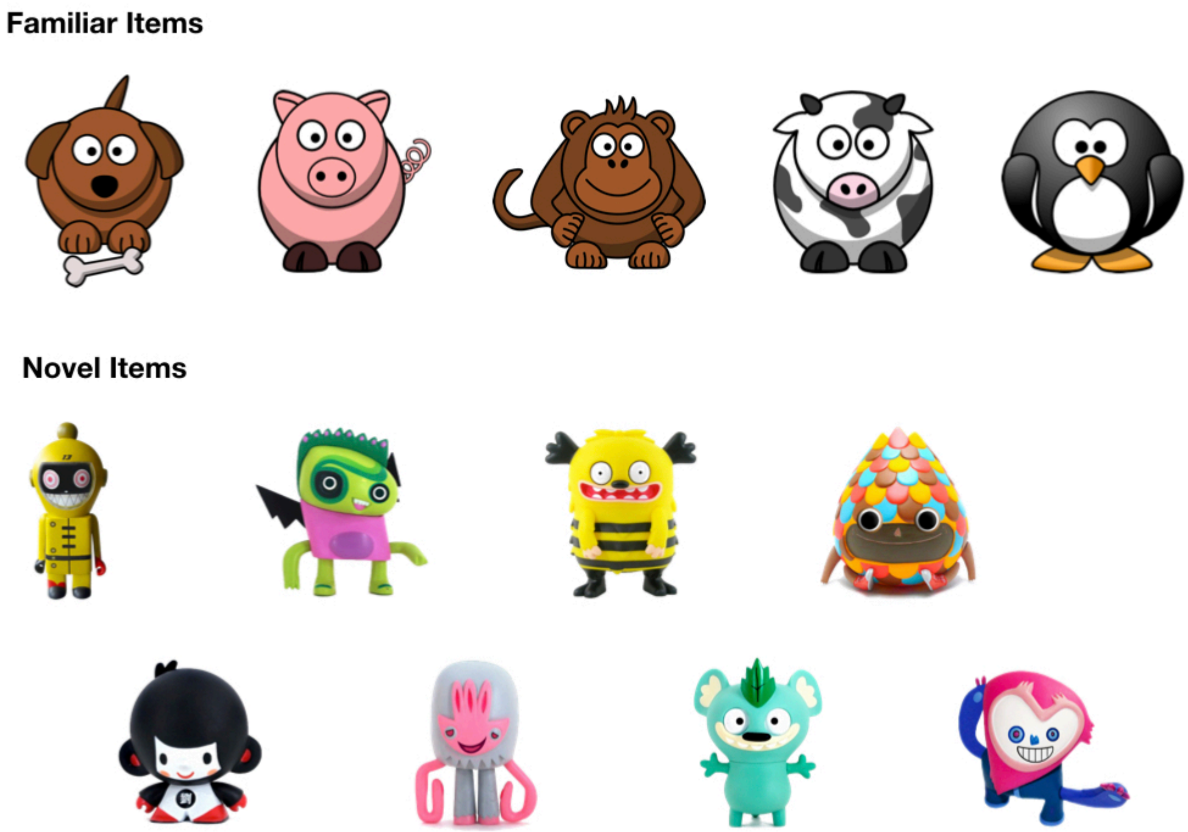

Figure A1. Visual stimuli for Experiments 2A and 2B in chapter 2. All objects drawn from a set used in previous word learning experiments with children (Partridge et al., 2015).

Figure A2. Visual stimuli for experiments in chapter 3. All objects drawn from a set used in previous word learning experiments with children (Partridge et al., 2015).

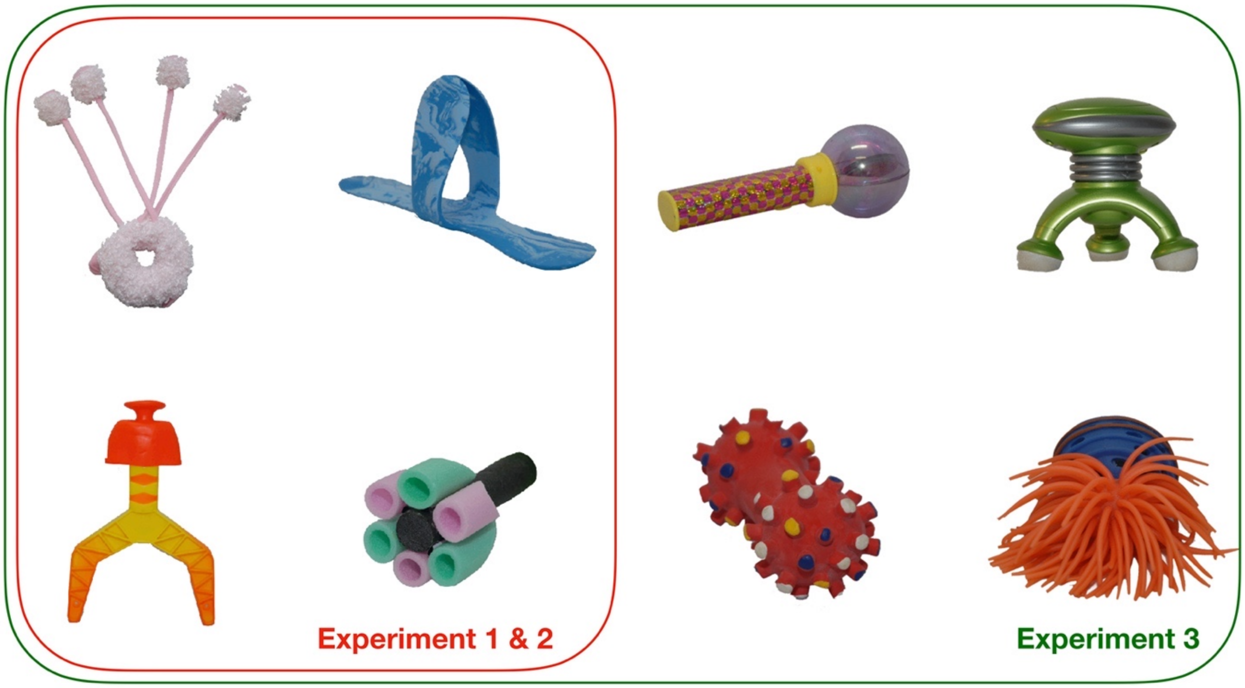

Figure A3. Visual stimuli for Experiments 1, 2, and 3 in chapter 4. All objects drawn from the Novel Object and Unusual Name (NOUN) Database (2nd edition) (Horst & Hout, 2016).

Figure A4. Visual stimuli for Experiments 1 and 2 in chapter 5. All objects drawn from the Novel Object and Unusual Name (NOUN) Database (2nd edition) (Horst & Hout, 2016).

Appendix B: Comparing accuracy for yoked pairings in the Active condition and the Yoked Passive Exposure Mismatch condition in chapter 3

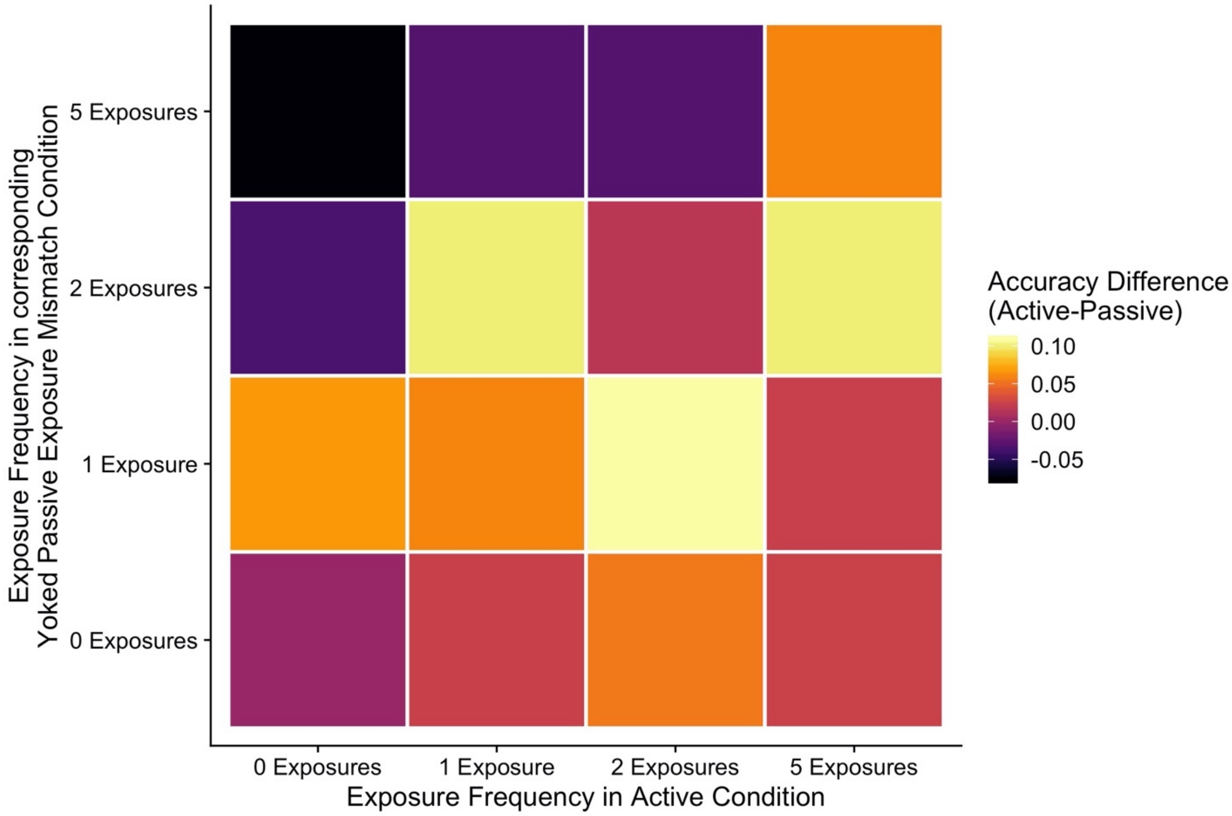

Figure B1. Differences in accuracy (Active – Passive) for yoked counterparts in the Active and Yoked Passive Exposure Mismatch condition, depending on exposure frequency. The x-axis corresponds to the frequency of exposure for a given item and participant in the Active condition. The y-axis corresponds the frequency of exposure for the corresponding item and yoked participant in the Yoked Passive Exposure Mismatch condition. The color fill represents the average difference in accuracy between the Active condition and the Yoked Passive Exposure Mismatch condition for a given item and yoked pairing. Darker colors represent a greater advantage for performance in the Yoked Passive Exposure Mismatch condition, and lighter/ yellower colors represent an advantage for performance in the Active condition. To test whether exposure difference predicted item-level accuracy difference, we fit a linear mixed-effects model predicting difference in test accuracy from difference in exposure frequency for a given item, including random intercepts for yoked pairing and item. Exposure difference was marginally related to difference in accuracy, $b = 0.01$, $F(1, 537.6) = 3.26$, $p = .07$. There was no interaction with Block.

References

Ackermann, L., Hepach, R., & Mani, N. (2019). Children learn words easier when they are interested in the category to which the word belongs. *Developmental Science*, 1–12. https://doi.org/10.1111/desc.12915

Adams, K. A., Kachergis, G., Gunzelmann, G., Howes, A., Tenbrink, T., & Davelaar, E. (2017). Executive function and attention predict low-income preschoolers' active category learning. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 57–62). Austin, TX: Cognitive Science Society.

Akhtar, N., & Tomasello, M. (2000). The social nature of words and word learning. In R. M. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. B. Smith, A. L. Woodward, N. Akhtar, … G. Hollich (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 115–135). Oxford: Oxford University Press.

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, *42*, 239–273. https://doi.org/10.1017/S030500091400049X

Aslin, R. N. (2014). Infant learning: Historical, conceptual, and methodological challenges. *Infancy*, *19*(1), 2–27. https://doi.org/10.1111/infa.12036

Baer, C., Gill, I. K., & Odic, D. (2018). A domain-general sense of confidence in children. *Open Mind*, *2*(2), 86–96. https://doi.org/10.1162/opmi_a_00020

Baer, C., & Odic, D. (2019). Certainty in numerical judgments develops independently of the approximate number system. *Cognitive Development*, *52*, 100817. https://doi.org/10.1016/j.cogdev.2019.100817

Barto, A., Mirolli, M., & Baldassarre, G. (2013). Novelty or surprise? *Frontiers in Psychology*, *4*, 1–15. https://doi.org/10.3389/fpsyg.2013.00907

Bates, D., & Maechler, M. (2009). *lme4: Linear mixed-effects models using S4 classes.* Retrieved from https://cran.r-project.org/web/packages/lme4/index.html

Bazhydai, M., Westermann, G., & Parise, E. (2020). "I don't know but I know who to ask": 12-month-olds actively seek information from knowledgeable adults. *Developmental Science*, 1–10. https://doi.org/10.1111/desc.12938

Begus, K., Gliga, T., & Southgate, V. (2014). Infants learn what they want to learn: Responding to infant pointing leads to superior learning. *PLoS ONE*, *9*(10), e108817. https://doi.org/10.1371/journal.pone.0108817

Begus, K., & Southgate, V. (2012). Infant pointing serves an interrogative function. *Developmental Science*, *15*(5), 611–617. https://doi.org/10.1111/j.1467-7687.2012.01160.x

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. New York, NY: McGraw-Hill.

Berlyne, D. E. (1966). Curiosity and exploration. *Science*, *153*(3731), 25–33.

Bloom, L. (2000a). The intentionality model of word learning: How to learn a word, any word. In R. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. Smith, A. Woodward, N. Akhtar, … G. Hollich (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 19–50). New York, NY: Oxford University Press.

Bloom, L. (2000b). What can we take for granted in word learning? In N. Akhtar, R. M. Golinkoff, K. Hirsh-Pasek, G. Hollich, L. B. Smith, M. Tomasello, & A. L. Woodward (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 165–168). New York, NY: Oxford University Press.

Bloom, L., Margulis, C., Tinker, E., & Fujita, N. (1996). Early conversations and word learning:

Contributions from child and adult. *Child Development*, *67*(6), 3154–3175.

Bloom, L., Tinker, E., & Scholnick, E. K. (2001). The intentionality model and language acquisition: Engagement, effort, and the essential tension in development. *Monographs of the Society for Research in Child Development*, *66*(4), i-viii+1-101. https://doi.org/10.4135/9781452286143.n496

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Bornstein, M. H., Putnick, D. L., Cote, L. R., Haynes, O. M., & Suwalsky, J. T. D. (2015). Mother-infant contingent vocalizations in 11 countries. *Psychological Science*, *26*(8), 1272–1284. https://doi.org/10.1177/0956797615586796

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2019). Consistency and variability in children's word learning across languages. *Open Mind*, *3*, 52–67. https://doi.org/10.1162/opmi_a_00026

Brennan, W. M., Ames, E. W., & Moore, R. W. (1966). Age differences in infants' attention to patterns of different complexities. *Science*, *151*(3708), 354–356.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, *31*, 21–32.

Butterworth, G. (2003). Pointing is the royal road to language for babies. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 9–33). Mahwah, NJ: Erlbaum.

Carvalho, P. F., Vales, C., Fausey, C. M., & Smith, L. B. (2018). Novel names extend for how long preschool children sample visual information. *Journal of Experimental Child Psychology*, *168*, 1–18. https://doi.org/10.1016/j.jecp.2017.12.002

Casillas, M., Brown, P., & Levinson, S. C. (2019). Early language experience in a Tseltal Mayan village. *Child Development*, 1–17. https://doi.org/10.1111/cdev.13349

Castro, R., Kalish, C., Nowak, R., Qian, R., Rogers, T., & Zhu, X. (2009). Human active learning. In D. Koller, D. Schuurmans, Y. Bengio, & L. Bottou (Eds.), *Advances in neural information processing systems* (pp. 241–249).

Chouinard, M. M. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development*, *72*(1), 1–129. https://doi.org/10.4135/9781452286143.n496

Clerkin, E. M., Hart, E., Rehg, J. M., Yu, C., & Smith, L. B. (2017). Real-world visual statistics and infants' first-learned object names. *Philosophical Transactions of the Royal Society B*, *372*, 20160055. https://doi.org/10.1098/rstb.2016.0055

Coenen, A., Nelson, J. D., & Gureckis, T. M. (2019). Asking the right questions about the psychology of human inquiry: Nine open challenges. *Psychonomic Bulletin and Review*, *26*, 1548–1587. https://doi.org/10.3758/s13423-018-1470-5

Coenen, A., Rehder, B., & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology*, *79*, 102–133. https://doi.org/10.1016/j.cogpsych.2015.02.004

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*, 337–367. https://doi.org/10.1146/annurev.psych.52.1.337

Cook, C., Goodman, N. D., & Schulz, L. E. (2011). Where science starts: Spontaneous experiments in preschoolers' exploratory play. *Cognition*, *120*(3), 341–349. https://doi.org/10.1016/j.cognition.2011.03.003

Dalmaijer, E. S., Mathôt, S., & Van der Stigchel, S. (2014). PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods*, *46*(4), 913–921. https://doi.org/10.3758/s13428-013-0422-2

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a

Web browser. *Behavior Research Methods*, *47*, 1–12. https://doi.org/10.3758/s13428-014-0458-y

Denison, S., Bonawitz, E., Gopnik, A., & Griffiths, T. L. (2013). Rational variability in children's causal inferences: The Sampling Hypothesis. *Cognition*, *126*(2), 285–300. https://doi.org/10.1016/j.cognition.2012.10.010

Destan, N., Hembacher, E., Ghetti, S., & Roebers, C. M. (2014). Early metacognitive abilities: The interplay of monitoring and control processes in 5- to 7-year-old children. *Journal of Experimental Child Psychology*, *126*, 213–228. https://doi.org/10.1016/j.jecp.2014.04.001

Feinman, S. (1982). Social referencing in infancy. *Merrill-Palmer Quarterly*, *28*(4), 445–470.

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics*, *21*, 95–115. https://doi.org/10.1017/S0142716400001053

Fiedler, K. (2008). The ultimate sampling dilemma in experience-based decision making. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(1), 186–203. https://doi.org/10.1037/0278-7393.34.1.186

Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science*, *20*(5), 578–585. https://doi.org/10.1111/j.1467-9280.2009.02335.x

Frank, M. C., Tenenbaum, J. B., & Fernald, A. (2013). Social and discourse contributions to the determination of reference in cross-situational word learning. *Language Learning and Development*, *9*(1), 1–24. https://doi.org/10.1080/15475441.2012.707101

Ghetti, S., Hembacher, E., & Coughlin, C. A. (2013). Feeling uncertain and acting on it during the preschool years: A metacognitive approach. *Child Development Perspectives*, *7*(3), 160–165. https://doi.org/10.1111/cdep.12035

Goldin-Meadow, S. (2007). Pointing sets the stage for learning language - and creating language. *Child Development*, *78*(3), 741–745. https://doi.org/10.1111/j.1467-8624.2007.01029.x

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology*, *64*(1), 257–283. https://doi.org/10.1146/annurev-psych-113011-143802

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences*, *100*(13), 8030–8035. https://doi.org/10.1073/pnas.1332441100

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*(5), 515–523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, *35*(3), 515–531. https://doi.org/10.1017/s0305000907008641

Gopnik, A., O'Grady, S., Lucas, C. G., Griffiths, T. L., Wente, A., Bridgers, S., … Dahl, R. E. (2017). Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proceedings of the National Academy of Sciences*, *114*(30), 7892–7899. https://doi.org/10.1073/pnas.1700811114

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, *138*(6), 1085–1108. https://doi.org/10.1037/a0028044

Gottlieb, J., Oudeyer, P.-Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends in Cognitive Sciences*, *17*(11), 585–593. https://doi.org/10.1016/j.tics.2013.09.001

Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they don't know. *Proceedings of the National Academy of Sciences*, *113*(13), 3492–3496. https://doi.org/10.1073/pnas.1515129113

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*, *30*(6), 509–516. https://doi.org/10.1177/0165025406071914

Gros-Louis, J., West, M. J., & King, A. P. (2014). Maternal responsiveness and the development of directed vocalizing in social interactions. *Infancy*, *19*(4), 385–408. https://doi.org/10.1111/infa.12054

Gros-Louis, J., & Wu, Z. (2012). Twelve-month-olds' vocal production during pointing in naturalistic interactions: Sensitivity to parents' attention and responses. *Infant Behavior and Development*, *35*(4), 773–778. https://doi.org/10.1016/j.infbeh.2012.07.016

Gureckis, T. M., & Markant, D. B. (2012). Self-directed learning: A cognitive and computational perspective. *Perspectives on Psychological Science*, *7*(5), 464–481. https://doi.org/10.1177/1745691612454304

Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology*, *53*(4), 310–344. https://doi.org/10.1016/j.cogpsych.2006.04.003

Heath, S. B. (1983). *Ways with words*. Cambridge: Cambridge University Press.

Held, R., & Hein, A. (1963). Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology*, *56*(5), 872–876.

https://doi.org/10.1037/h0040546

Hembacher, E., & Frank, M. C. (2017). Children's social referencing reflects sensitivity to graded uncertainty. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 495–500). Austin, TX: Cognitive Science Society.

Henderson, J. M. (2017). Gaze control as prediction. *Trends in Cognitive Sciences*, *21*(1), 15–23. https://doi.org/10.1016/j.tics.2016.11.003

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, *33*(2–3), 61–83; discussion 83-135. https://doi.org/10.1017/S0140525X0999152X

Hidaka, S., Torii, T., & Kachergis, G. (2017). Quantifying the impact of active choice in word learning. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 519–525). Austion, TX: Cognitive Science Society.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*(1), 55–88. https://doi.org/10.1016/j.dr.2005.11.002

Hollich, G., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., … Bloom, L. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, *65*(3), 1–135. https://doi.org/10.4135/9781452286143.n496

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

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*(2), 128–157. https://doi.org/10.1080/15250000701795598

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

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248. https://doi.org/10.1037/0012-1649.27.2.236

Itti, L., & Baldi, P. (2006). Bayesian surprise attracts human attention. *Advances in Neural Information Processing Systems*, *19*, 1295–1306. https://doi.org/10.1016/j.visres.2008.09.007

Jiminez, S., Sun, Y., & Saylor, M. M. (2018). The process of active word learning. In M. M. Saylor & P. Ganea (Eds.), *Active Learning from Infancy to Childhood: Social Motivation, Cognition, and Linguistic Mechanisms* (pp. 75–93). Cham, Switzerland: Springer.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Kachergis, G., Rhodes, M., & Gureckis, T. (2017). Desirable difficulties during the development of active inquiry skills. *Cognition*, *166*, 407–417. https://doi.org/10.1016/j.cognition.2017.05.021

Kachergis, G., Yu, C., & Shiffrin, R. M. (2012). An associative model of adaptive inference for learning word-referent mappings. *Psychonomic Bulletin and Review*, *19*(2), 317–324.

https://doi.org/10.3758/s13423-011-0194-6

Kachergis, G., Yu, C., & Shiffrin, R. M. (2013). Actively learning object names across ambiguous situations. *Topics in Cognitive Science*, *5*(1), 200–213. https://doi.org/10.1111/tops.12008

Kachergis, G., Yu, C., & Shiffrin, R. M. (2014). Cross-situational word learning is both implicit and strategic. *Frontiers in Psychology*, *5*, 1–10. https://doi.org/10.3389/fpsyg.2014.00588

Keijser, D., Gelderloos, L., & Alishahi, A. (2019). Curious topics: A curiosity-based model of first language word learning. In *Proceedings of the 41st Annual Conference of the Cognitive Science Society* (pp. 1991–1997). Montreal, QB: Cognitive Science Society.

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron*, *88*, 449–460. https://doi.org/10.1016/j.neuron.2015.09.010

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS One*, *7*(5), e36399. https://doi.org/10.1371/journal.pone.003639

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks effect in infant auditory attention. *Child Development*, *85*(5), 1795–1804. https://doi.org/10.1111/cdev.12263

Klinnert, M. D., Emde, R. N., Butterfield, P., & Campos, J. J. (1986). Social referencing: The infant's use of emotional signals from a friendly adult with mother present. *Developmental Psychology*, *22*(4), 427–432. https://doi.org/10.1037/0012-1649.22.4.427

Kovács, Á. M., Tauzin, T., Téglás, E., Gergely, G., & Csibra, G. (2014). Pointing as epistemic request: 12-month-olds point to receive new information. *Infancy*, *19*(6), 543–557. https://doi.org/10.1111/infa.12060

Kurkul, K. E., & Corriveau, K. H. (2018). Question, explanation, follow-up: A mechanism for learning from others? *Child Development*, *89*(1), 280–294. https://doi.org/10.1111/cdev.12726

Legare, C. H. (2012). Exploring explanation: Explaining inconsistent evidence informs exploratory, hypothesis-testing behavior in young children. *Child Development*, *83*(1), 173–185. https://doi.org/10.1111/j.1467-8624.2011.01691.x

Legare, C. H., Gelman, S. A., & Wellman, H. M. (2010). Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child Development*, *81*(3), 929–944. https://doi.org/10.1111/j.1467-8624.2010.01443.x

Legare, C. H., Mills, C. M., Souza, A. L., Plummer, L. E., & Yasskin, R. (2013). The use of questions as problem-solving strategies during early childhood. *Journal of Experimental Child Psychology*, *114*(1), 63–76. https://doi.org/10.1016/j.jecp.2012.07.002

Lewis, M., Cristiano, V., Lake, B. M., Kwan, T., & Frank, M. C. (2020). The role of developmental change and linguistic experience in the mutual exclusivity effect. *Cognition*, *198*, 104191. https://doi.org/10.1016/j.cognition.2020.104191

Lipko, A. R., Dunlosky, J., Hartwig, M. K., Rawson, K. A., Swan, K., & Cook, D. (2009). Using standards to improve middle school students' accuracy at evaluating the quality of their recall. *Journal of Experimental Psychology: Applied*, *15*(4), 307–318. https://doi.org/10.1037/a0017599

Lipko, A. R., Dunlosky, J., Lipowski, S. L., & Merriman, W. E. (2012). Young children are not underconfident With practice: The benefit of ignoring a fallible memory heuristic. *Journal of Cognition and Development*, *13*(2), 174–188. https://doi.org/10.1080/15248372.2011.577760

Lipowski, S. L., Merriman, W. E., & Dunlosky, J. (2013). Preschoolers can make highly accurate judgments of learning. *Developmental Psychology*, *49*(8), 1505–1516. https://doi.org/10.1037/a0030614

Lucca, K., & Wilbourn, M. P. (2018a). Communicating to learn: Infants' pointing gestures result in optimal learning. *Child Development*, *89*(3), 941–960. https://doi.org/10.1111/cdev.12707

Lucca, K., & Wilbourn, M. P. (2018b). The what and the how: Information-seeking pointing gestures facilitate learning labels and functions. *Journal of Experimental Child Psychology*, *178*, 417–436. https://doi.org/10.1016/j.jecp.2018.08.003

Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development*, *82*(6), 1778–1787. https://doi.org/10.1111/j.1467-8624.2011.01649.x

MacDonald, K., & Frank, M. C. (2016). When does passive learning improve the effectiveness of active learning? In *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 2459–2464). Austin, TX: Cognitive Science Society.

Mani, N., & Ackermann, L. (2018). Why do children learn the words they do? *Child Development Perspectives*, *12*(4), 253–257. https://doi.org/10.1111/cdep.12295

Markant, D. B., DuBrow, S., Davachi, L., & Gureckis, T. M. (2014). Deconstructing the effect of self-directed study on episodic memory. *Memory and Cognition*, *42*(8), 1211–1224. https://doi.org/10.3758/s13421-014-0435-9

Markant, D. B., & Gureckis, T. M. (2014). Is it better to select or to receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology. General*, *143*(1), 94–122. https://doi.org/10.1037/a0032108

Markant, D. B., Ruggeri, A., Gureckis, T. M., & Xu, F. (2016). Enhanced memory as a common effect of active learning. *Mind, Brain, and Education*, *10*(3), 142–152. https://doi.org/10.1111/mbe.12117

Markant, D. B., Settles, B., & Gureckis, T. M. (2016). Self-directed learning favors local, rather than global, uncertainty. *Cognitive Science*, *40*(1), 100–120. https://doi.org/10.1111/cogs.12220

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

Matatyaho, D. J., & Gogate, L. J. (2008). Type of maternal object motion during synchronous naming predicts preverbal infants' learning of word–object relations. *Infancy*, *13*(2), 172–184. https://doi.org/10.1080/15250000701795655

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(22), 9014–9019. https://doi.org/10.1073/pnas.1105040108

Mills, C. M., Legare, C. H., Grant, M. G., & Landrum, A. R. (2011). Determining who to question, what to ask, and how much information to ask for: The development of inquiry in young children. *Journal of Experimental Child Psychology*, *110*(4), 539–560. https://doi.org/10.1016/j.jecp.2011.06.003

Miyazaki, M., Takahashi, H., Rolf, M., Okada, H., & Omori, T. (2014). The image-scratch paradigm: A new paradigm for evaluating infants' motivated gaze control. *Scientific Reports*, *4*, 1–6. https://doi.org/10.1038/srep05498

Morey, R. D. (2008). Confidence Intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64.

https://doi.org/10.20982/tqmp.04.2.p061

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends in Cognitive Sciences*, *15*(10), 453–459. https://doi.org/10.1016/j.tics.2011.07.011

Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current Directions in Psychological Science*, *21*(2), 71–77. https://doi.org/10.1177/0963721412436807

Nelson, J. D., Divjak, B., Gudmundsdottir, G., Martignon, L. F., & Meder, B. (2014). Children's sequential information search is sensitive to environmental probabilities. *Cognition*, *130*(1), 74–80. https://doi.org/10.1016/j.cognition.2013.09.007

Oudeyer, P.-Y., Kachergis, G., & Schueller, W. (2019). *Computational and robotic models of early language development: A review*. Retrieved from http://arxiv.org/abs/1903.10246

Oudeyer, P.-Y., & Smith, L. B. (2016). How evolution may work through curiosity-driven developmental process. *Topics in Cognitive Science*, *8*(2), 492–502. https://doi.org/10.1111/tops.12196

Partridge, E., Mcgovern, M. G., Yung, A., & Kidd, C. (2015). Young children's self-directed information gathering on touchscreens. In R. Dale, C. Jennings, P. Maglio, T. Matlock, D. Noelle, A. Warlaumont, & J. Yoshimi (Eds.), *Proceedings of the Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pelz, M., Piantadosi, S. T., & Kidd, C. (2015). The dynamics of idealized attention in complex learning environments. *5th Joint International Conference on Development and Learning* *and Epigenetic Robotics, ICDL-EpiRob 2015*, 236–241. https://doi.org/10.1109/DEVLRN.2015.7346147

Piaget, J. (1952). *The origins of intelligence in children*. New York: International University Press.

Piaget, J. (1955). *The child's construction of reality*. London: Routledge & Kegan Paul.

Piaget, J. (1964). Development and learning. In R. E. Ripple & V. N. Rockcastle (Eds.), *Piaget Rediscovered* (pp. 7–20).

Piaget, J. (1976). *Piaget sampler: An introduction to Jean Piaget through his own words*. (S. F. Campbell, Ed.). New York: Wiley.

Plato. (1997). *Complete Works*. (J. M. Cooper, Ed.). Indianapolis: Hackett.

Pomper, R., & Saffran, J. R. (2019). Familiar object salience affects novel word learning. *Child Development*, *90*(2), e246–e262. https://doi.org/10.1111/cdev.13053

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

R Development Core Team. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Roembke, T. C., & McMurray, B. (2016). Observational word learning: Beyond propose-but-verify and associative bean counting. *Journal of Memory and Language*, *87*, 105–127. https://doi.org/10.1016/j.jml.2015.09.005

Roembke, T. C., & McMurray, B. (2020). Interactions of explicit and implicit learning mechanisms in cross-situational word learning. PsyArXiv. https://doi.org/10.31234/osf.io/3k7s8

Romberg, A. R., & Yu, C. (2014). Interactions between statistical aggregation and hypothesis testing mechanisms during word learning. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Meeting of the Cognitive Science* (pp. 1311–1316). Austin, TX: Cognitive Science Society.

Romberg, A. R., & Yu, C. (2015). Interactions between statistical aggregation and hypothesis testing mechanisms during word learning. In E. Grillo & K. Jepson (Eds.), *Proceedings of the 39th Annual Boston University Conference on Language Development* (pp. 1311–1316). Somerville, MA: Cascadilla Press.

Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in early gesture explain. *Science*, *323*(13), 951–953. https://doi.org/10.1126/science.1167025

Rowe, M. L., Özçalişkan, Ş., & Goldin-Meadow, S. (2008). Learning words by hand: Gesture's role in predicting vocabulary development. *First Language*, *28*(2), 182–199. https://doi.org/10.1177/0142723707088310

Roy, B. C., Frank, M. C., Decamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *PNAS*, *112*(41), 12663–12668. https://doi.org/10.1073/pnas.1419773112

Ruff, H. A., & Capozzoli, M. C. (2003). Development of attention and distractibility in the first 4 years of life. *Developmental Psychology*, *39*(5), 877–890. https://doi.org/10.1037/0012-1649.39.5.877

Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition*, *143*, 203–216. https://doi.org/10.1016/j.cognition.2015.07.004

Ruggeri, A., Lombrozo, T., Griffiths, T. L., & Xu, F. (2016). Sources of developmental change in the efficiency of information search. *Developmental Psychology*, *52*(12), 2159–2173. https://doi.org/10.1037/dev0000240

Samuelson, L., & Smith, L. (1998). Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, and Tomasello. *Child Development*, *69*(1), 94–104. https://doi.org/10.1111/j.1467-8624.1998.tb06136.x

Saylor, M. M., & Ganea, P. A. (2018). *Active Learning from Infancy to Childhood*. Cham, Switzerland: Springer.

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, *43*(4), 1045–1050. https://doi.org/10.1037/0012-1649.43.4.1045

Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science*, *10*(3), 322–332. https://doi.org/10.1111/j.1467-7687.2007.00587.x

Settles, B. (2012). *Active Learning*. *Synthesis Lectures on Artificial Intelligence and Machine Learning* (Vol. 6). https://doi.org/10.2200/S00429ED1V01Y201207AIM018

Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science*, *7*(4), 341–351. https://doi.org/10.1177/1745691612448481

Shanks, D. R., Tunney, R. J., & McCarthy, J. D. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making*, *15*(3), 233–250. https://doi.org/10.1002/bdm.413

Sim, Z. L., Tanner, M., & Alpert, N. Y. (2015). Children learn better when they select their own data. In P. P. Noelle, D. C., Dale, R., Warlaumont, A. S., Yoshimi, J., Matlock, T., Jennings, C. D., & Maglio (Ed.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 2194–2199). Austin, TX: Cognitive Science Society.

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences*, *22*(4), 325–336. https://doi.org/10.1016/j.tics.2018.02.004

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

Snyder, H. R., & Munakata, Y. (2010). Becoming self-directed: Abstract representations support endogenous flexibility in children. *Cognition*, *116*(2), 155–167. https://doi.org/10.1016/j.cognition.2010.04.007

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, *348*(6230), 91–94. https://doi.org/10.1126/science.aaa3799

Stahl, A. E., & Feigenson, L. (2019). Violations of core knowledge shape early learning. *Topics in Cognitive Science*, *11*(1), 136–153. https://doi.org/10.1111/tops.12389

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

The ManyBabies Consortium. (2020). Quantifying sources of variability in infancy research using the infant-directed speech preference. *Advances in Methods and Practices in Psychological Science*. https://doi.org/10.1177/2515245919900809

Tomasello, M., & Akhtar, N. (1995). Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cognitive Development*, *10*(2), 201–224.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *The Behavioral and Brain Sciences*, *28*(5), 675–691; discussion 691-735. https://doi.org/10.1017/S0140525X05000129

Trueswell, J. C., Lin, Y., Armstrong, B., Cartmill, E. A., Goldin-Meadow, S., & Gleitman, L. R.

(2016). Perceiving referential intent: Dynamics of reference in natural parent-child interactions. *Cognition*, *148*, 117–135. https://doi.org/10.1016/j.cognition.2015.11.002

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, *66*(1), 126–156. https://doi.org/10.1016/j.cogpsych.2012.10.001

Tsuji, S., Jincho, N., Mazuka, R., & Cristia, A. (2020). Communicative cues in the absence of a human interaction partner enhance 12-month-old infants' word learning. *Journal of Experimental Child Psychology*, *191*, 104740. https://doi.org/10.1016/j.jecp.2019.104740

Twomey, K. E., & Westermann, G. (2017). Curiosity-based learning in infants: A neurocomputational approach. *Developmental Science*, 1–13. https://doi.org/10.1111/desc.12629

Vaish, A., Demir, Ö. E., & Baldwin, D. (2011). Thirteen- and 18-month-old infants recognize when they need referential information. *Social Development*, *20*(3), 431–449. https://doi.org/10.1111/j.1467-9507.2010.00601.x

van Schijndel, T. J. P., Visser, I., van Bers, B. M. C. W., & Raijmakers, M. E. J. (2014). Preschoolers perform more informative experiments after observing theory-violating evidence. *Journal of Experimental Child Psychology*, *131*, 104–119. https://doi.org/10.1016/j.jecp.2014.11.008

Vouloumanos, A., & Werker, J. F. (2009). Infants' learning of novel words in a stochastic environment. *Developmental Psychology*, *45*(6), 1611–1617. https://doi.org/10.1037/a0016134

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Walden, T. A., & Ogan, T. A. (1988). The development of social referencing. *Child Development*, *59*(5), 1230–1240. https://doi.org/10.1111/j.1467-8624.1988.tb01492.x

Wang, Q., Bolhuis, J., Rothkopf, C. A., Kolling, T., Knopf, M., & Triesch, J. (2012). Infants in control: Rapid anticipation of action outcomes in a gaze-contingent paradigm. *PloS One*, *7*(2), e30884. https://doi.org/10.1371/journal.pone.0030884

Wass, Sam, Porayska-Pomsta, K., & Johnson, M. H. (2011). Training attentional control in infancy. *Current Biology*, *21*(18), 1543–1547. https://doi.org/10.1016/j.cub.2011.08.004

Wass, S V, Smith, T. J., & Johnson, M. H. (2013). Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. *Behavior Research Methods*, *45*(1), 229–250. https://doi.org/10.3758/s13428-012-0245-6

Wu, Z., & Gros-Louis, J. (2014). Infants' prelinguistic communicative acts and maternal responses: Relations to linguistic development. *First Language*, *34*(1), 72–90. https://doi.org/10.1177/0142723714521925

Wu, Z., & Gros-Louis, J. (2015). Caregivers provide more labeling responses to infants' pointing than to infants' object-directed vocalizations. *Journal of Child Language*, *42*(3), 538–561. https://doi.org/10.1017/s0305000914000221

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414–420. https://doi.org/10.1111/j.1467-9280.2007.01915.x

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262. https://doi.org/10.1016/j.cognition.2012.06.016

Yu, Y., Shafto, P., Bonawitz, E., Yang, S. C. H., Golinkoff, R. M., Corriveau, K. H., … Xu, F. (2018). The theoretical and methodological opportunities afforded by guided play with young children. *Frontiers in Psychology*, *9*, 1–8. https://doi.org/10.3389/fpsyg.2018.01152

Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition*, *145*, 53–62. https://doi.org/10.1016/j.cognition.2015.07.013

Yurovsky, D., & Frank, M. C. (2017). Beyond naive cue combination: Salience and social cues in early word learning. *Developmental Science*, *20*(2), e12349. https://doi.org/10.1111/desc.12349

Zipf, G. (1965). *Human behavior and the principle of least effort: An introduction to human ecology.* New York: Hafner.