THE UNIVERSITY OF CHICAGO

LEARNING TO TALK: THE DEVELOPMENT OF CHILDREN'S COMMUNICATIVE SKILLS

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

BY

ASHLEY CHUI-KAY LEUNG

CHICAGO, ILLINOIS

AUGUST 2023

Table of Contents

List of Figures iv

Acknowledgments v

Abstract viii

Introduction 1

Chapter 1. Parents adapt speech to their young children's vocabulary knowledge 14

1.1. Experiment 1.1: Testing the fine-tuning hypothesis in parents' speech to their toddlers 15

1.1.1. Methods 15

1.1.2. Results 18

1.1.3. Discussion 25

Chapter 2. Parents scaffold the formation of conversational pacts with their children 28

2.1. Experiment 2.1: Forming pacts in an interactive communication game 32

2.1.1. Methods 34

2.1.2. Results and Discussion 36

2.2. Experiment 2.2: Comprehension 44

2.2.1. Methods 45

2.2.2. Results 47

2.2.3. Discussion 49

2.3. Experiment 2.3: Production 50

2.3.1. Methods 52

2.3.2. Results 56

2.3.3. Discussion 62

2.4. General Discussion 64

Chapter 3. A Closer Look at Communicative Perspective-Taking: Children's Estimates of Other's Specific Knowledge 67

3.1. Experiment 3.1: Children's estimates of other's vocabulary knowledge 72

3.1.1. Methods 72

3.1.2. Results 76

3.1.3. Discussion 82

3.2. Experiment 3.2: Do children spontaneously adapt to a conversational partner? 84

3.2.1. Methods 85

3.2.2. Preliminary Results and Discussion 87

Chapter 4. What facilitates communicative perspective-taking? An online computation account of communication 89

References 119

List of Figures

Figure 1.1 A parent-child playing the interactive game in Experiment 1.1 16

Figure 1.2 Parents' length of expressions reflect beliefs about child's knowledge 23

Figure 1.3 Parents modulate length of expression based on children's accuracy 23

Figure 1.4 Graphical representation of qualitative analysis of Experiment 1.1 24

Figure 2.1 Schematic of experimental design and trial structure of Experiment 2.1 33

Figure 2.2 Accuracy and number of dialogue exchanges per trial 34

Figure 2.3 Number of words produced across rounds 38

Figure 2.4 Analysis of the source of referential pacts 41

Figure 2.5 Results for Experiment 2.2 48

Figure 2.6 Schematic of trials in Experiment 2.3 52

Figure 2.7 Context sensitivity and nameability in both adults and children 57

Figure 2.8 Performance on ambiguous manipulation check trials predicts context-sensitivity ... 61

Figure 3.1 Trial structure for Experiment 3.1 74

Figure 3.2 Comparing adult AoA estimates to children's judgments 78

Figure 3.3 Children's judgments of word knowledge across development 78

Figure 3.4 Coefficient estimates of the effect of AoA on participant judgments 80

Figure 3.5 Children's explanations 82

Figure 3.6 Example trials for Experiment 3.2 86

Acknowledgments

This dissertation is a culmination of a lot of interesting questions, painstakingly-designed experiments (that sometimes still didn't work), and trying to make sense of data by plotting everything 20 ways for visualization. More importantly, it is the product of an incredibly supportive group of advisors, co-workers, family, and friends. Each person involved in my academic and personal achievements in the last 6 years has meant so much to me, and it would be impossible to thank all of you, but I will try to do so here.

First of all, I want to thank my academic advisors. Dan, you are an incredibly understanding, honest, and humble advisor. I'm honored to be one of your first students, and it's bittersweet to know that I'm also one of your last. I am so glad I got to work with you these last 6 years, and I'm very proud of the projects we've done together. Susan, thank you for taking me under your wing for the past 3 or so years. I admired your work from afar before coming to UChicago, and I've gotten to admire it up close in the last 6 years. I'm in awe of the passion you have for research, and the joy with which you share about your work. Alex, your classes are some my favorites from grad school. I don't know if it's the brutal honesty, or the fast talking, but you always generate great conversation from the group. Thank you also for being a caring mentor—whether I was a student in your class, or a TA for your section in Mind, I always knew that you had my best interests in mind and cared for my wellbeing and success. Middy, one of my biggest regrets is that our timing didn't quite match up for us to work together on any projects. You started around the time everything locked down, and I moved away soon after things went back in person. I admire your research, your teaching, and the way you mentor and lead. Thank you for being a mentor to me in both my academic and personal life. Robert, you started out as a collaborator and became a mentor and a friend. Your wisdom, attention to detail, and willingness to make time for the people your mentor speaks volumes about you as a researcher and a person. Thank you for all the guidance you've given me in my research and my career.

To my lab mates, you are the most wonderful coworkers and friends I could have ever asked for. Claire, I've been blown away by your intelligence, your boldness, and your many, many (did I mention many?) talents. Thank you for fun conversations about everything under the sun. I'm so glad we got to start and finish our PhD's together, and I can't wait to see what you do next. I know you'll do incredible things, and I'll be cheering you on wherever we both are. Ben, you're one of the kindest souls I know. Thank you for being a wonderful friend, an amazing collaborator, and a great sounding board for all things academic and life. The three of us have been through quite the journey, and as cliché as it sounds, there's nobody else I'd rather have weathered the storms with. May there be many more conference trips ahead of us.

To the friends and church family I met when I was in Evanston—you have been my rock in so many ways. It's hard to live away from biological family, and I'm so lucky to have had all of you to lean on. From game nights to shared meals and deep conversations, I've thoroughly loved living life with all of you. I want to thank everyone we met at Evanston Bible Fellowship, but especially: Charlotte Abrahamson, Michael Shi, Chelsea Sherlock, Moses Chan, Jacinda Lee, Tina and Antonio Chan Gonzalez, and Tom and Ann Marie Hartmann. Your friendship and discipleship have shaped me in so many ways. I treasure the memories we've shared together and look forward to the new ones we will build. To the community that I'm now building in the Boston area, thank you for taking us in with welcoming arms. I especially want to thank my Community Group at Aletheia—you have made our transition to this new city a lot smoother, and I look forward to growing together with you all.

I would not be here (literally) if not for my family. Mom and Dad, thank you for encouraging me to pursue my interests, even if that meant I would move halfway across the world. Mom, thank you for being my confidante, my friend, and my devil's advocate. I admire how you're always learning new things. Not many people can say their mom studied for the SATs with them, and I'm so lucky to have someone who did that for me. Dad, you're the one who inspired me to pursue this path, and I'm so glad to have someone like you to look up to. Thank you for instilling in me a sense of curiosity, a passion for travel, and an adventurous palate. You have both formed and shaped me into the person I am today. Anna is luck to have you as her grandparents. To Kerry, thank you for sharing a room with me until I left for college, even though we had an extra bedroom. You had to put up with a fair bit of temper and drama from me, and I can honestly say your (and Mom's) patience and calm is something I can only hope to have one day. I'm so glad you're at least in the same country, so I never have to go too long without seeing you.

Finally, the people who have become the most important to me. Edmund, aside from my family, you've known me the longest. You've seen the literal worst and best of me, and you've loved me through it all. There are too many things to thank you for, so I'll choose one here: thank you for believing in me, no matter how little I believe in myself. I often joke that I take a lot of credit for your PhD, but it goes both ways—this is as much my achievement as it is yours. To Anna, it will be a while before you know how to read this, but I've wanted to be a mom for as long as I can remember, and you made that dream come true. Your joy, curiosity, and excitement are so refreshing, and I hope you never lose that spark. Being your mom will forever be one of my proudest achievements.

Soli Deo Gloria

Abstract

Children learn language at incredible rates, and a plethora of research has examined the mechanisms and trajectories of young children's language development. However, human language consists of more than words and syntax—it conveys meaning for the purpose of communicating with others. This dissertation centers on the following question: How do children learn to use language to communicate? To begin tackling this question, I consider the kinds of interactions that children engage in. First, I ask how parents play a role in scaffolding children's communicative development by adapting the way they speak, both generally and specifically, to their children's developmental level and knowledge (Chapters 1 and 2). I find that parents of children (ages 2-8) adjust the vocabulary as well as communicative strategies they use during cooperative games to ensure communicative success. Then, I ask how children might learn to communicate with peers, and the development of component skills required for adapting to others in conversation (Chapter 3). I find that while children as young as 4 are sensitive to what words are learned earlier or later (even when they themselves know all the words in question), they may not spontaneously adapt their language when communicating with another child. Finally, I present a theoretical framework of conversational adaptation, specifically considering how perspective-taking occurs (or fails to occur) in both adults and children. I hypothesize that in-the-moment processing demands is the key factor driving communicative success and failure in adults and children alike (Chapter 4). This dissertation seeks to understand the development of children's communicative skill by examining the interactive contexts that children engage in, and aims to present a holistic picture of language learning—one that does not isolate the child from their environment, but one that considers the rich input that children receive through interactions that are designed for them.

Introduction

Children learn language at a surprisingly rapid rate, knowing thousands of words by the time they are toddlers. But language development goes beyond learning words and syntax, or even meaning. Language is interactive by nature, and learning language is learning to *communicate*. A wealth of prior research shows that caregiver input predicts children's language development (e.g., Hart & Risley, 1995; Huttenlocher et al., 1991; Huttenlocher et al., 2010). Importantly, input during *interactions*, where language is directed to children, and where children engage in conversational turns, is a particularly powerful predictor for language learning outcomes (Romeo et al., 2018; Weisleder & Fernald, 2013). Even so, the majority of research on language development focus on specific aspects or characteristics of children's linguistic input, and not the actual interactive contexts that children are in. In my dissertation, I aim to capture the rich communicative interactions that children engage in to uncover why certain environments are particularly helpful for children's communicative development. The overarching question that I will seek to address in this dissertation is: How do children learn to communicate with others?

One key component of communication is adaptation—to understand one another, conversational partners consider what the other person knows, and adjust accordingly. While adults adapt to others in conversation with relative ease, children often struggle to do so on a variety of referential tasks (e.g., Krauss & Glucksberg, 1977). One difficulty that children may face is making appropriate inferences about their partner's knowledge—if they are unable to estimate another person's knowledge, they may not see the need to make adaptations in communication. Alternatively, children may be sensitive to their conversational partners' knowledge, but struggle to leverage that information during online communication. To date, many studies have attempted to ask *when* children learn to communicate effectively with others, but few studies have investigated *how* children might gain this ability.

In my dissertation, I will consider children's communicative development in interactive settings. First, what role might parent-child interactions play in facilitating the development of children's communicative abilities? Given the plethora of research showing the importance of caregiver input on language development, studying the interactions between parents and children is likely to shed light onto the ways in which these contexts support broader communicative development. Second, what might be some reasons for children's apparent struggle to adapt to others? In this part of my dissertation, I will investigate children's ability to make inferences about other's knowledge, and to subsequently leverage those inferences to adapt to others in online communication. Taken together, the goal of my dissertation is to understand the contexts which may support children's communicative development, as well as the developmental trajectory of children's communicative skills.

The Role of Parental Input in Children's Language Development

From birth (and even *in utero*), children are exposed to the language of their caregivers. Shortly after birth, children already recognize and prefer their mother's voices and their native language (DeCasper & Fifer, 1980; Moon et al., 1993). As children develop, their caregivers continue to play a crucial role in shaping their language learning. The quantity and quality of caregiver input have been shown to predict children's language development, with more recent works highlighting the importance of socially contingent language interactions (e.g., Romeo et al., 2018). Why is parental input a potentially powerful predictor of children's language learning?

Even before children can produce language or participate in spoken conversation, their environments are full of linguistic input. Of all the linguistic content that children are exposed to, some are directed *at them*. That is, language spoken directly to chcildren, or meant for children to hear. Many studies have concluded that it is *child-directed speech*, and not merely overheard speech, that predicts children's language learning (e.g., Huttenlocher et al., 1991; Weisleder & Fernald, 2013). In a pivotal study by Hart and Risley (1995), researchers found that children from lower social economic status (SES) backgrounds receive reduced language input from their caregivers, potentially explaining later differences in children's language development and other academic outcomes. The quantity and diversity of caregivers' speech to children significantly predicts toddlers' vocabulary, and mediates the effect of SES on language development, highlighting the importance of parental input on children's learning (Huttenlocher et al., 2010). Weisleder and Fernald (2013) also found that greater amounts of child-directed speech correlated with expressive vocabulary development, mediated by language processing efficiency. In other words, children who encounter more child-directed speech may be more efficient at processing language, which in turn stimulates vocabulary growth.

The aforementioned studies point to the power of child-directed speech and appear to suggest that *quantity* of children's linguistic input correlates with language learning outcomes. What about the *quality* of language input? The above studies focus on the amount of speech produced by parents during parent-child interactions—how important is the *interactive* nature of language for facilitating children's communicative development? Romeo et al. (2018) found that 4- to 6-year-old children's language exposure, specifically *conversational turns* between children and their parents, most significantly predicted language skills. This effect was mediated by differences in neural structures in children's brains, suggesting that engaging in more conversational turns may lead to structural changes in certain subregions of the brain's language area, which then facilitates language development. Importantly, Romeo et al.'s (2018) work illustrates that child-directed speech may be powerful not simply because it is directed at the child, but because it often involves participation from the child.

Language is interactive in nature, and language learning is necessarily more than just acquiring vocabulary and grammar—learning language is learning to *communicate*. It might be unsurprising, therefore, that contingency is a key feature of child-directed speech. Starting from infancy, children receive linguistic input that is adapted to them. Caregivers' speech to 9-month-olds are simplified when produced within 2 seconds of infants' babbles, suggesting that parents may be engaging in proto-conversation by adapting the way they speak in response to their infants' vocalizations (Elmlinger, Schwade, & Goldstein, 2019). Parents also produce certain words more frequently around the time their children acquire those specific words, indicating incredible sensitivity to children's vocabulary development (Roy, Frank, & Roy, 2009). These studies suggest that parents are sensitive and *responsive* to their young children's emerging language skills.

How might parental responsiveness, or *adaptation*, facilitate children's language development? It is possible that parents can leverage their extensive knowledge of their children's developmental level to provide input that is optimal for them, thus *scaffolding* language learning (Vygotsky, 1978). The *linguistic tuning hypothesis*, proposed by Snow (1972), posits that child-directed speech is helpful for language learning because of its adaptive nature. That is, parents adjust their language as their children develop, ensuring that input is tuned to an appropriate level for learning. In an observational study, Masur (1992) found that parents talk differently about objects that their children know versus those their children do not. For example, parents ask questions to elicit children's labelling when talking about toys that their children are familiar with, but are more likely to label unfamiliar toys themselves (Masur, 1992). In Chapter 1 of this dissertation, I provide the first *empirical test* of the linguistic tuning hypothesis, and consider how parents tune speech at both coarse and fine levels to their toddlers during an interactive communication game.

Communication as Adaptation

Thus far I have presented evidence that children's language input is adapted to them, and that that adaptive nature of the language they hear may be helpful for their language learning. In this section, I venture one step further to consider communication in general as necessarily adaptive. That is, for anyone to communicate successfully with one another, adaptations must be made. The extent to which adjustments are made will depend on a variety of factors, such as the distance between two conversational partners' knowledge, linguistic abilities, and so on.

Communication is a collaborative, dynamic process involving at least two conversational partners. One line of thinking considers communication as coordination. Clark (1996) argues that when people converse with one another, they engage in joint actions, much as two people aiming to complete a manual task together. As individuals coordinate with one another, they can activate any shared knowledge, or common ground, to facilitate their communication. Many studies have investigated how common ground is built and modified over conversation, showing that adults seamlessly form and modify common ground (e.g., Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986). Even when referring to common objects, conversational partners engage in negotiation to form conceptual pacts, or temporary agreements on referent names (Brennan & Clark, 1996). For example, when referring to a shoe when multiple types are present, interlocuters may use names such as "loafer" to uniquely identify their intended referent. This conceptual pact persists even when the other types of shoes are removed, but not when a new conversational partner is introduced. Brennan and Clark interpret these findings to mean that conceptual pacts are partner-specific and serve to facilitate efficient communication between partners. The conceptualization of communication as *adaptation* follows from the idea that communication is a form of coordination.

Coordination requires some form of understanding another person's perspective, whether through estimation or through explicit questioning. Thus, successful coordination requires adaptation. Conversational partners can adjust the way they speak at the level of syntax, or even select specific words, in order to most effectively communicate with their partner. We can imagine, however, that adaptation varies across communicative settings. When speaking with a familiar peer, for example, we may need relatively few adjustments due to extensive shared knowledge, as well as relatively matched linguistic and cognitive abilities. On the other hand, when speaking with a young child that we meet for the first time, we would likely adapt more significantly to the child, such as by using simpler words, or speak about topics that likely interest children.

A large body of work has investigated communicative perspective-taking in adults, and to a lesser extent, in children. This line of work often utilizes paradigms where two conversational partners differ in visual perspective—some objects in the environment are visible to one partner but not to the other. By creating asymmetries in visual perspective, researchers can ask how accurately and rapidly people consider their partner's perspectives during referential communication. These studies show that even adults, under certain circumstances, commit errors in perspective-taking. However, the explanations for these errors differ widely. One account claims that considering other perspectives is effortful, and errors in perspective-taking arise from reverting to an egocentric "default" when there are time pressures, cognitive load, etc. (e.g., Keysar et al., 2000). A second account argues that we weight multiple perspectives incrementally and dynamically (Hanna et al., 2003), while a third account posits that conversational partners rationally consider the "division of labor" between partners based on repeated interactions (e.g., Hawkins et al., 2021). I present each of these frameworks in detail in Chapter 4, and discuss the empirical findings associated with each account. While the specific findings and conclusions drawn by researchers differs, this body of work highlights the importance of perspective-taking, or adaptation, in communication—the effectiveness of our daily conversations rely upon our ability to adapt to one another.

Communicative Development

While specific experimental manipulations can lead to higher error rates in even adults' perspective-taking, adults generally adapt to one another with relative ease (e.g., Clark & Wilkes-Gibbs, 1986). However, many children struggle to adapt to peers in similar referential settings (e.g., Krauss & Glucksberg, 1977). In Krauss & Glucksberg's studies, even older children (9^{th} graders) had difficulty communicating about abstract shapes to their peers. Although some studies have shown that children can form referential pacts with each another in some situations, they do not do so reliably until middle childhood (e.g., Branigan et al., 2016). Importantly, adult feedback may play a role—the participants in Branigan et al.'s study were told whether they communicated successfully (i.e., whether the correct target object was selected by a conversational partner). Further, there is evidence suggesting that explicit adaptation from adults may facilitate successful referential communication. Krauss et al. (1966) found that even preschool-aged children could select the correct target when child-generated names for abstract referents were used by an experimenter. This finding indicates that one challenge children may face is in recognizing that another person's perspective differs from theirs.

Prior to interacting with peers, children's early communicative exposures likely involve adults, often their caregivers. What roles might caregivers play in facilitating children's ability to adapt to others? While parents may not always provide the same kinds of explicit feedback as experimenters in Branigan et al.'s (2016) study, they may nonetheless provide scaffolding for their children's communicative development. Parents, rather than adults in general, may be particularly able to provide relevant support to their children due to familiarity about their specific children's developmental level. In Chapter 2 of this dissertation, I ask how parents and children form referential pacts with each other. I ask whether children, who struggle to communicate with peers, can successfully form referential pacts with the support of a caregiver. What strategies to parents use to facilitate referential success, and how does parent speech change over children's development? In the same chapter, I also investigate children's comprehension and production skills when removed from a scaffolded context, to explore the possible challenges leading to children's struggles when communicating with peers.

While parental scaffolding can provide a supportive context for young children's early communicative development, there are also significant linguistic and cognitive skills that children acquire that likely support communication. For example, the development of children's executive functioning and theory of mind is related to their communicative perspective-taking ability (Nilsen & Graham, 2009; Sidera et al., 2018). Preschoolers with greater inhibitory control exhibit less egocentric processing, suggesting that the ability to consider other perspectives may partially arise from an ability to inhibit one's own perspective (Nilsen & Graham, 2009). Similarly, children's theory of mind skills, or their ability to hold multiple perspectives in mind, are important for communicative success. Sidera et al found that 6- to 10-year-old children's performance on a battery of theory of mind measures correlated with their cooperative success in a referential task. Taken together, these studies suggest that children's communicative development is facilitated, at least in part, by their cognitive development.

What other skills must children have before they can adapt to their conversational partners effectively? One important ability may be estimating others' knowledge. A crucial limitation in the current literature is that many studies ask how children overcome visual perspective differences in communication, but not how they address difference in knowledge states. Unlike differences in visual perspectives, mismatches in knowledge between individuals are not usually observable. How does sensitivity to differences in knowledge emerge in childhood, and how does this ability impact children's communicative development?

Prior research shows that young children have sensible intuitions about people's general knowledge, and can attribute different levels of knowledge to infants, preschool children, and adults (Fitneva, 2010; Taylor, et al., 1991). For example, by age 5, children recognize that an adults and preschoolers, but not an infant, knows what an elephant is (Taylor et al., 1991). Children as young as 3 years old also distinguish between child-specific and adult-specific information and recognize that there are domains where children may be more knowledgeable than adults, such as information related cartoon characters or toys (Fitneva, 2010; VanderBorght & Jaswal, 2009). Young children also prefer to learn from people who they believe to be more knowledgeable, indicating that children's inferences about others have important consequences for their subsequent interactions with them (VanderBorght & Jaswal).

While children appear to be adept at estimating the general knowledge of others, very little is known about children's inferences about other people's specific knowledge. Why is this an important skill? Communication relies not only on general adaptation at the discourse level (e.g., choosing to talk about toys with children), but also on specific adaptations at the word level (e.g., using the word *bird* instead of *peacock* when speaking to a toddler). Some studies investigating children's general knowledge inferences also include vocabulary items (e.g., Taylor et al., 1991), but only test whether children can make broad distinctions (e.g., that a child knows *happy* but not *hypochondriac*). Children in Taylor et al.'s study could have used their own word knowledge to make judgments. That is, they may infer that other children also know words that they themselves know (e.g., *happy*). Thus, an open question remains as to whether children can make finer-grained estimates of another person's vocabulary knowledge.

Interestingly, one study found that children have relatively accurate estimates of their *own* word learning (Walley & Metsala, 1992). Walley and Metsala found that children as young as age 5 reliably recovered the order in which they learned various lexical items, and even made reasonable estimates for when they might learn certain words in the future. While this study did not ask children to make inferences about other children's vocabulary knowledge, the results indicate that children have surprising metalinguistic knowledge. If children can maintain fine-grained details and estimates of their own word learning, can they do the same for someone else? Some research shows that children may over-attribute knowledge (e.g., Birch & Bloom, 2003), and it is thus possible that children would over-estimate the vocabulary of a younger child. In Chapter 3 of this dissertation, I ask whether young children (ages 4-8) have accurate estimates of other people's specific vocabulary knowledge. In an exploratory follow-up experiment, I ask whether children's models of other people's knowledge has consequences for communicative adaptation.

Theories of Communicative Adaptation

Currently, multiple frameworks for understanding communicative adaptations in adults has been proposed, but very little developmental work has engaged with these theories. Why is a theoretical framework that includes children important? First, children engage in conversations, much like adults. While children may lack some of the verbal and cognitive skills that adults have, the processes by which communication occurs is likely to differ greatly between adults and children. A comprehensive theory of communication should therefore include developmental perspectives. Second, children provide an interesting opportunity to understand the factors that drive our communicative success, precisely because they are "poorer" conversationalists. By studying situations where children may struggle to take the perspectives of others, we may clarify and quantify the factors the ultimately drive our communicative adaptations.

Communication is often successful, but even adults occasionally make errors in perspective-taking. A common thread in the current literature has been to investigate these errors as a way to shed light on the communicative process. In Chapter 4, I present discuss three existing frameworks of communicative adaptation (or communicative perspective-taking): two-step models, probabilistic models, and division-of-labor models (Hanna et al., 2003; Hawkins et al., 2021; Keysar et al., 2000). Each of these models provides potential explanations for successful and unsuccessful perspective taking, and each lend support from certain empirical studies. However, none of these frameworks engage directly with developmental data, nor do they claim to make predictions about children.

Therefore, in Chapter 4, I also propose a novel framework: an *online computation account*, which posits that the ease of communicative perspective-taking changes with the amount of *in-the-moment computation* that speakers and listeners are required to do. Much of the information that we leverage in communication may actually be pre-computed, and thus easily accessed during online conversation, and these pre-computed models that we have of various conversational partners can lend itself to rapid conversational adaptation. The goal of my proposed framework is to provide a theory that has explanatory and predictive power for both adult and developmental empirical studies.

Overview of Chapters

In Chapter 1, I present the first empirical test of the linguistic tuning hypothesis. Using an interactive communication game, I find that parents *fine-tune* their speech to their toddlers (ages 2-2.5) in service of effective communication. When parents talk about animals that they believe their children know, they use shorter descriptions. On the other hand, when talking about animals they believe their children do not know, parents use longer descriptions, suggesting that they are providing additional information to facilitate children's understanding. Further, I show that parents can also adjust their beliefs in real time, producing longer sentences when describing an animal that the child previously failed to recognize during the game. Finally, I find that parents use various strategies to communicate effectively with their children: they use more adjectives, more comparisons to other labels, and more basic level category labels when talking about animals that they think their children do not know.

In Chapter 2, I ask how parents scaffold children's (ages 4-8) referential abilities in a matcher-director task. I find that children as young as 4 can form referential pacts when interacting with their parents, and that parent speech during the cooperative communication task changes as a function of children's age. Specifically, parents talk more to younger children, and ask more questions to younger children. These findings suggest that parents adapt to their children's developmental level, and provide a supportive context for children as they learn to navigate the referential process. In two follow-up experiments, I and ask how children's comprehension or production skills impact their referential abilities. I find that children do not appear to have trouble comprehending referential expressions, but may have difficulty generating informative descriptions of referents.

In Chapter 3, I ask whether children (ages 4-8) have accurate representations of other people's specific vocabulary knowledge. Children were introduced to a younger fictional child and asked to estimate that child's vocabulary. I find that even 4-year-olds make judgments that are in line with adult-reported estimates of age of acquisition (AoA). That is, children accurately infer that a younger child is likely to know words that are generally learned earlier (e.g., *dog*), and less likely to know words that are often learned later (e.g., *lobster*). There is also significant developmental change, such that 8-year-olds' judgments of the younger child's vocabulary knowledge more reliably reflected adult-reported AoA. Further, children largely appeal to linguistic reasons when asked why they made certain inferences about the younger child's knowledge, suggesting that metalinguistic knowledge plays a role in children's ability to infer others' lexical knowledge. In an ongoing follow-up experiment, children (ages 4-8) are asked to communicate with a younger fictional child. Based on data from a partial sample, I find that children do not appear to modify their descriptions based on beliefs about the younger child's knowledge. That is, despite accurately representing that a younger child may not know certain words (e.g., *penguin*), they do not adapt their speech accordingly.

Finally, in Chapter 4, I present three existing theories of communicative adaptation, and evaluate their claims against adult and developmental data. I point out that a common missing link is that current theories lack direct predictions and explanations for developmental findings, and propose a comprehensive framework: an online computation account of communicative adaptation. I discuss how the proposed framework may explain the variance across both adult and developmental research, as well as its relevance to Chapters 1-3 of this dissertation.

Chapter 1. Parents adapt speech to their young children's vocabulary knowledge

In just a few short years, children master their native language. The quantity and quality of caregivers' linguistic input, especially input that occurs during interactions, are powerful predictors of children's language development (Hart & Risley, 1995; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Romeo et al., 2018; Rowe, 2012). Why is parental input so important to language learning?

The way we speak to children is markedly different from the way we speak to adults. Child-directed speech differs from adult-directed speech on various dimensions, such as acoustics, prosody, syntactic structure, and content (e.g., Cooper & Aslin, 1990; Newport, Gleitman, & Gleitman, 1977; Snow, 1972). Crucially, child-directed speech changes over development, with sentences getting longer, more complex, and less acoustically variable (e.g., Huttenlocher et al., 2010; Liu, Tsao, & Kuhl, 2009; Phillips, 1973). The linguistic tuning hypothesis suggests that this changing nature of child-directed speech is what allows it to be such a powerful driver of language development (Snow, 1972). If parents tune their speech to children's developmental level, increasing the complexity of input at the same rate that children are developing their linguistic knowledge, input may always be at the optimal level of complexity to support language learning (Vygotsky, 1978).

How precisely do parents tune their speech? One possibility is that caregivers coarse-tune: adjusting complexity of their speech generally using a holistic sense of their children's developing linguistic abilities. Over and above this coarse-tuning, parents might fine-tune their speech, taking into account not only children's global linguistic development, but their specific knowledge of smaller units of language, such as lexical items. Fine-tuning would provide a particularly powerful and efficient vehicle for scaffolding language acquisition because of its specificity. To date, the only evidence for fine-tuning comes from observational studies (Masur, 1997; Roy, Frank, & Roy, 2009). Here, we present the first experimental evidence for fine-tuning.

Children and their parents played a reference game in which the parent's goal was to guide their child to select a target animal from a set of three. Parents tuned the amount of information in their utterances not just to the average difficulty of each animal word, but to their prior estimates of their individual child's knowledge of that animal. Further, parents sensitively adapted over the course of the game, providing more information on subsequent trials when they discovered that their child did not know an animal. Together, these results show that parents leverage their knowledge of their children's language development to fine-tune the linguistic information they provide.

1.1. Experiment 1.1: Testing the fine-tuning hypothesis in parents' speech to their toddlers

1.1.1. Methods

Participants

Toddlers (aged 2-2.5 years) and their parents were recruited to achieve a planned sample of 40 parent-child dyads. A total of 48 parent-child pairs were recruited, but data from 7 pairs were dropped from analysis because of failure to complete the experiment as designed. The final sample consisted of 41 children aged 24 mo.; 5 days to 29 mo.; 20 days ($M = 26$ mo.; 0 days), 21 of whom were girls. A total of 48 parent-child pairs were recruited, but data from 7 pairs were dropped from analysis because of failure to complete the experiment as designed. Of the 7 pairs that were dropped, 5 children fussed out, 1 had an older sibling interfering with the study, and 1 was a twin (only the twin who participated first was included). The final sample consisted of 41 children aged 24 mo.; 5 days to 29 mo.; 20 days ($M = 26$ mo.; 0 days), 21 of whom were girls.

Stimuli

Eighteen animal images were selected from the Rossion and Pourtois (2004) image set, a colorized version of the Snodgrass and Vanderwart (1980) object set. Animals were selected based on estimates of their age of acquisition (AoA) for American English learners, using information from parent reports on Wordbank (Frank, Braginsky, Yurovsky, & Marchman, 2017) and retrospective self-report estimates from adults (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). Half of the animals were chosen to have an early AoA (15-23 months), and the other half were chosen to have a late AoA (25-32 months). Each trial featured three animals, all from either the early AoA or late AoA category.

A modified version of the MacArthur-Bates Communicative Development Inventory Short Form (CDI; Fenson et al., 2007), a parent-reported measure of children's vocabulary, was administered before the testing session via an online survey. The selected animal words were added to the standard words, producing an 85-word survey. Two of the animal words–one in the early AoA (*pig*) and one in the late AoA category (*rooster*)–were accidentally omitted, so trials for those words were not included in analyses as we could not obtain individual-level estimates of children's knowledge.

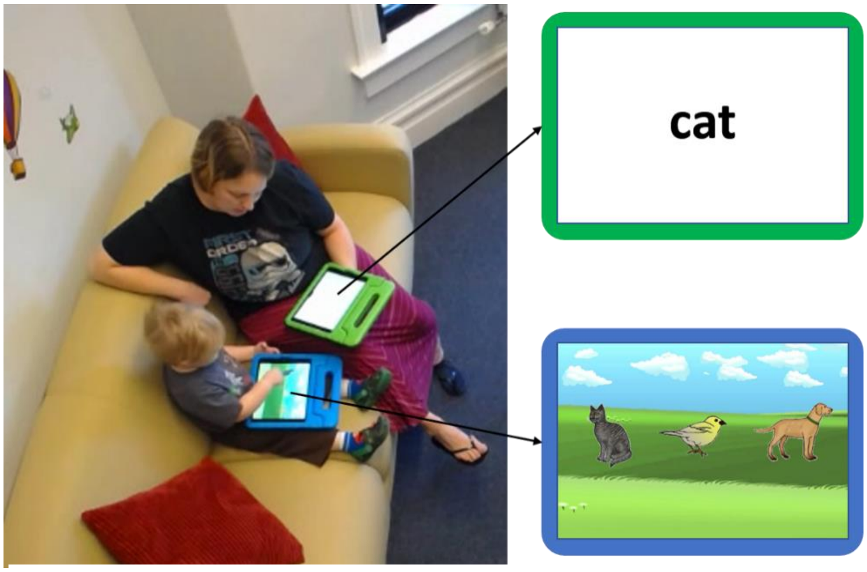

Figure 1.1 *A parent-child pair playing the interactive game in Experiment 1.1*

Design and Procedure

Each parent-child pair played an interactive reference game using two iPads (***Figure 1.1***). Children began with two warm-up trials in which they tapped on circles that appeared on the iPads. Following these warm-up trials, children and their parents moved on to practice and then experimental trials. On each trial, three images of animals were displayed side by side on the child's screen, and a single word appeared on the parent's screen. Parents were instructed to communicate as they normally would with their child, and to encourage their child to choose the object corresponding to the word on their screen. The child was instructed to listen to their parent for cues. Once the child tapped on an animal, the trial ended and a new trial began. There were 36 experimental trials, such that each animal appeared as the target twice. Trials were randomized for each participant, with the constraint that the same animal could not be the target twice in a row. Practice trials followed the same format as experimental trials, with the exception that images of fruit and vegetables were shown. All sessions were videotaped for transcription and coding.

Data Analysis

Our primary quantity of interest was the amount of information that parents provided in each of their utterances. To approximate this, we measured the length of parents' referring expressions–the number of words they produced on each trial before their child selected an animal. Length is an imperfect proxy for information, but it is easy to quantify and theory-agnostic. Because utterance length is highly right-skewed (i.e., most utterances are short), we log-transformed length in all analyses. However, to facilitate interpretability, we show raw utterance length in our figures. Subsequently, utterances were manually coded for the following: (1) Use of an animal's canonical label (e.g., "leopard"), (2) Use of a descriptor (e.g., "spotted"), (3) Use of a comparison (e.g., "like a cat"), (4) Use of a superordinate level category label (e.g., "bird" for peacock), and (5) Use of a subordinate level category label (e.g., "Limelight Larry," a fictional character from a children's book, for peacock). Parent utterances irrelevant to the game (e.g., asking the child to sit down) were not analyzed. Children's utterances were coded when audible, but were not analyzed. Our second source of data was the vocabulary questionnaire that parents filled out prior to participation. Parents indicated whether their child produced each of the 85 words on the survey. In addition to analyzing parents' judgments for the animals in the task, we also computed the total number of words judged to be known for each child as a proxy for total vocabulary.

All analyses were done using mixed-effects models. In all cases we began with maximal random effects structures and pruned random effects until the models converged. We removed interaction terms before removing main effects, and opted to keep the most theory-relevant random effects when only a subset of main effects could be kept.

1.1.2. Results

We begin by confirming that our *a priori* divisions of animals into early and late age of acquisition (AoA) in the study design were reflected in parents' survey judgments, and that children were able to follow parents' references to select the correct target animal on each trial. After this, we show that parents fine-tune their referring expressions, producing more information in their references to animals that they think their individual children do not know. Further, parents updated their tuning over the course of the experiment, producing more information on subsequent references to animals they thought their children knew after observing evidence to the contrary (i.e., children made an incorrect selection).

Target animal difficulty

We first confirm that the early AoA animals were more likely to be marked "known" by the parents of children in our studies. As predicted, parents judged that their children knew 94% of the animals in the early AoA category, and 33% of the animals in the late AoA category, which were reliably different from each other (β = -6.48, p < .001, d = -3.57 [-4.48, -2.67]).

Selection accuracy

On the whole, parents communicated effectively with their children, such that children selected the correct target on 69.05% of trials, reliably greater than would be expected by chance (33%, β = 2.07, p < .001, d = 1.14 [0.93, 1.36]). Children were above chance both for animals that parents thought they knew (M = 75.08%, β = 2.61, p < .001, d = 1.44 [1.18, 1.70]), and for animals that parents thought their children did not know (M = 55.19%, β = 1.23, p < .001, d = 0.68 [0.52, 0.84]). Thus, parents successfully communicated the target referent to children, even when parents thought their children did not know the name for the animal at the start of the game.

Was this accuracy driven by children's knowledge or parents' referring expressions? Because we did not measure children's knowledge of each animal directly, we used parents' estimates of children's knowledge as a proxy to answer this question. We fit a mixed-effects logistic regression predicting children's accuracy on each trial from children's total estimated vocabulary, parent-reported knowledge of the target animal, and the (log) length of parents' expressions. We found that children with bigger vocabularies were more accurate in general (β = 0.40, p = .001, d = 0.22 [0.09, 0.36]), and that children were less accurate for animals whose names parents thought they did not know (β = -1.86, p < .001, d = -1.02 [-1.46, -0.58]). Longer referring expressions were associated with greater accuracy for animals that parents thought their children did not know (β = 0.46, p = .025, d = 0.25 [0.03, 0.47]), but not for animals that parents thought their children knew (β = -0.40, p = .007, d = -0.22 [-0.38, -0.06]).

Thus, longer referring expressions were associated with more successful communication for animals that parents thought their children did not know, but not for animals that parents thought they knew. This suggests that parental tuning contributed to children's success. We next ask whether parents tuned the lengths of their utterances appropriately, producing longer expressions for animals they believe their children do not know.

Tuning

If parents calibrate their referring expressions to their children's linguistic knowledge, they should provide more information to children for whom a simple bare noun (e.g., "leopard") would be insufficient to identify the target. Parents did this in a number of ways: with one or more adjectives (e.g., "the spotted, yellow leopard"), with similes (e.g., "the one that's like a cat"), and with allusions to familiar animal exemplars of the category (e.g., "Limelight Larry," a fictional peacock from a children's book). In many of these cases, parents would be required to produce more words (see below for further qualitative analyses). Thus, we first analyzed the (log) length of parents' referring expressions as a proxy for informativeness.

When do parents produce longer referring expressions? One possibility is that parents tune at the coarsest level, using more words when speaking to children with smaller vocabularies. This was not the case–the total number of words parents thought their children knew did not reliably affect the length of their referring expressions (β = -0.02, p = .595, d = -0.17 [-0.79, 0.45]). A second possibility is that parents have a sense for how difficult each animal is in general, and tune coarsely to this. Our analyses confirmed this coarse-tuning: parents said reliably fewer words for animals that more children were reported to know (β = -0.17, p = .034, d = -1.19 [-2.26, -0.09]);

Figure 1.3). Finally, parents could fine-tune their referring expressions to their children's individual knowledge, over and above the average difficulty of each animal. Our analyses supported this conclusion: parents used reliably fewer words to refer to animals that they thought their individual child knew ($\beta = 0.25$, $p = .003$, d = 0.98 [0.34, 1.61]); **Figure 1.3**). Thus, parents fine-tuned the amount of information in their referring expressions, calibrating the amount of information they provided to their children's knowledge, even after accounting for the average difficulty of the target animal.

In addition, because each animal appeared as a target twice, we asked whether parents tuned their referring expressions over successive appearances. We found that parents used fewer words on the second appearance of each animal ($\beta =$ -0.08, $p = .044$, d = -1.06 [-2.07, -0.03]). This effect is shown in **Figure 1.3**: the mean length of referring expression is lower on the second appearance when compared to the first (**Figure 1.3**). However, the difference in utterance length between animals they thought their children knew versus didn't know was smaller on their second appearance ($\beta =$ -0.14, $p < .001$, d = -0.17 [-0.22, -0.12]; **Figure 1.3**). Why might that be? One possibility is that parents obtain information from the first appearance of each animal: they may have thought their child knew "leopard," but discovered from their incorrect choice that they did not. If so, they might provide more information the second time around.

To test this prediction, we fit a model predicting the (log) length of parents' referring expressions from appearance type (first, following correct, following incorrect), whether the parent thought their child knew the animal prior to the experiment, and their interaction between appearance type and prior belief. Relative to their utterances on an animal's first appearance, parents produced shorter referring expressions on an animal's second appearance following both correct responses ($\beta =$ -0.14, $p = .036$, d = -0.12 [-0.23, -0.01]) and incorrect responses ($\beta =$ -

0.28, $p = < .001$, d = -0.22 [-0.34, -0.11]). As before, parents produced shorter utterances for animals they thought their child knew ($\beta =$ -0.31, $p = < .001$, d = -0.92 [-1.43, -0.41]). When children were correct on an animal's first appearance, parents' referring expressions on its second appearance did not differ in length based on whether they thought their child knew the animal prior to the experiment ($\beta =$ -0.02, $p = .771$, d = -0.02 [-0.13, 0.10]). However, when children were incorrect on an animal's first appearance, and parents thought they knew the animal prior to the experiment, they produced reliably longer referring expressions on its second appearance ($\beta =$ 0.43, $p = < .001$, d = 0.24 [0.13, 0.35]; **Figure 1.3**).

As we predicted, when parents thought their children knew an animal, but then observed evidence to the contrary, they provided more information in their referring expressions to help children make the correct selection the second time. However, we did not find the opposite pattern: when children were successful for animals that parents thought they did not know, parents did not update their beliefs. Why should parents update their beliefs in one direction but not the other? One likely explanation comes from parents' linguistic tuning itself. Parents' goal in this task is to produce a referring expression that allows their children to select the target animal whether or not they know its canonical label. Consequently, when children select correctly on these trials, parents cannot know how their child arrived at the correct target. For example, if a parent said "spotted leopard" and their child was correct, they could not know whether their child succeeded because they actually knew the word "leopard," or instead because the word "spotted" was sufficiently informative.

Together, these two sets of analyses suggest that parents tune their referring expressions not just coarsely to how much language their children generally know, nor their knowledge about how hard animal words are on average, but finely to their beliefs about their individual children's knowledge of specific lexical items. Further, when parents discover that they have incorrect beliefs about their children's knowledge of an animal, they update these beliefs in real-time and adjust subsequent references to the same animal.

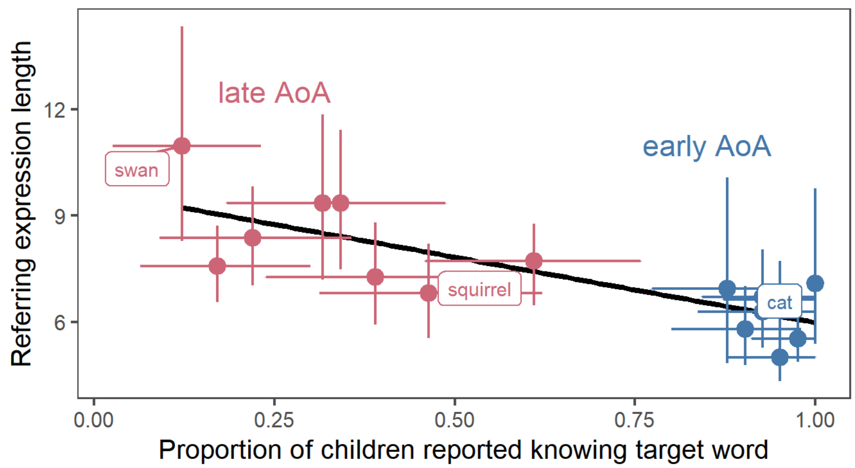

Figure 1.2 *Parents produced longer referring expressions to communicate about animals that children were generally less likely to know.*

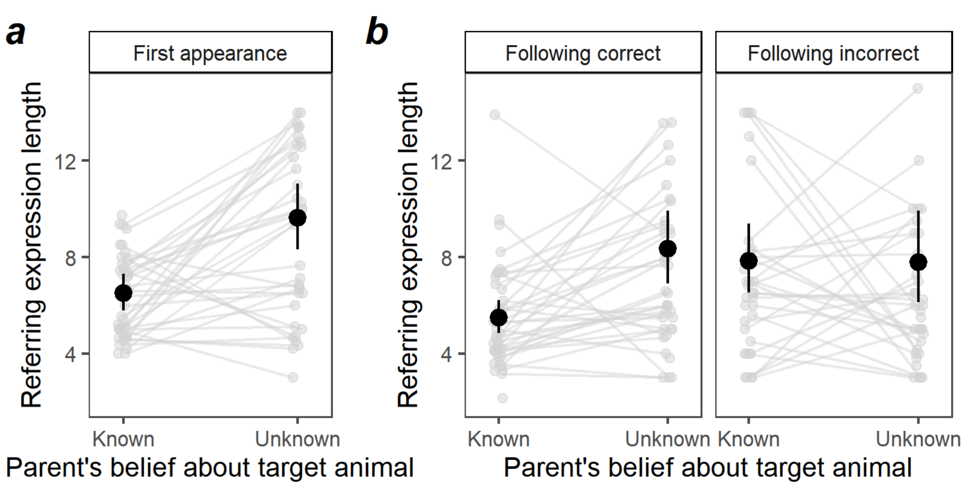

Figure 1.3 *(A) Parents produced longer referring expressions for animals they thought their child did not know (first appearance). (B) When children chose the correct animal, parents continued to produce longer expressions for animals they thought their children did not know (left). However, if parents thought their child knew an animal, and they chose incorrectly, parents produced longer expressions on its second appearance (right).*

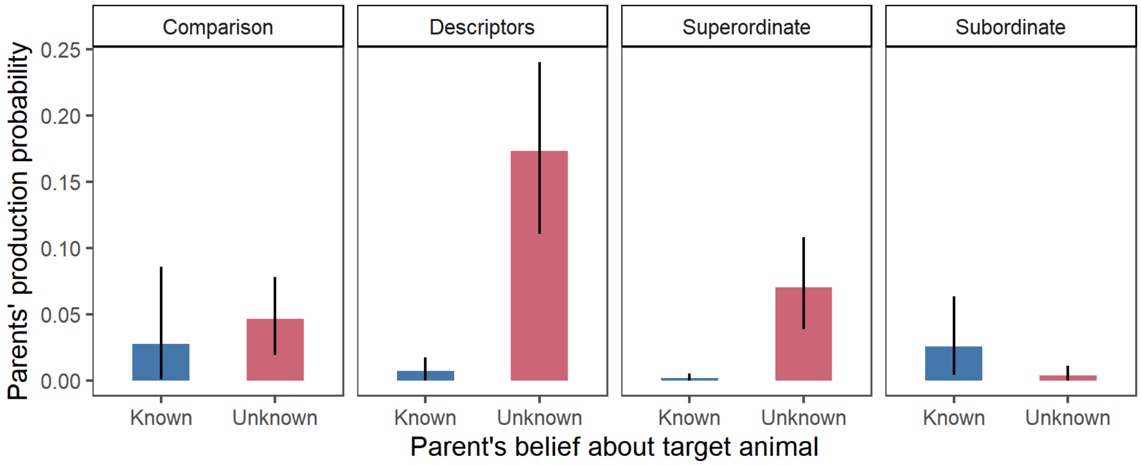

Figure 1.4 *Proportion of trials on which parents used comparison, descriptors, superordinate category labels, and subordinate category labels in their referring expressions.*

Content of referring expressions

Parents produced reliably longer referring expressions when trying to communicate about animals that they thought their children did not know. *How* did parents successfully refer to animals that their children did not know? As a post-hoc descriptive analysis, five qualitative features of referring expressions were coded: (1) Use of the animal's canonical label (e.g., "leopard"), (2) Use of a descriptor (e.g., "spotted"), (3) Use of a comparison (e.g., "like a cat"), (4) Use of a superordinate level category label (e.g., "bird" for peacock), and (5) Use of a subordinate level category label (e.g., "Limelight Larry" for peacock). Because the rates of usage of each of these kinds of reference varied widely (e.g. canonical labels were used on 94.48% of trials, but subordinates were used on 1.21% of trials), we fit a logistic mixed effects model separately for each reference kind, estimating whether it would be used on each trial from whether the parent thought their child knew the animal.

Canonical labels were used on almost all trials, and did not differ in frequency between animals parents thought their children did not know ($M = 94.68\%$) and animals they thought their children knew (M = 93.43%, β = 0.43, p = .216, d = 0.23 [-0.14, 0.61]). Parents thus produced canonical labels even when they thought their children did not know these labels. One plausible explanation for this is that the target animal on each trial was identified in writing for the parent, activating the canonical label and thus lowering the cost of retrieving and producing it (Wingfield, 1968). Another possibility is that this reflects parents' general tendency to produce basic-level category labels when talking to children (e.g., Blewitt, 1983). Finally, it could have been produced for implicitly or explicitly pedagogical reasons even though it was not referentially necessary. We expand on this possibility in the Discussion below.

Comparisons were used reliably more for animals parents believed their children did not know than for animals they thought their children knew (β = 2.29, p = .001, d = 1.26 [0.49, 2.04]), as were descriptors (β = 3.09, p < .001, d = 1.71 [1.07, 2.35]) and superordinate category labels (β = 3.01, p = .026, d = 1.66 [0.20, 3.12]). Subordinates were used less for animals parents thought their children did not know than for animals they thought their children knew (β = -2.19, p = .025, d = -1.21 [-2.26, -0.15]). Thus, parents used a variety of strategies to refer to animals that they believed their children do not know, but the use of descriptors was the most prominent (**Figure 1.4**). These descriptors are particularly apt to facilitate children's learning, connecting parents' fine-tuning for reference with their children's language acquisition.

1.1.3. Discussion

Parents have a wealth of knowledge about their kids, including about their linguistic development (Fenson et al., 2007). In this study, we asked whether parents leverage this knowledge to communicate successfully with their children. When playing a referential communication game, parents drew on their knowledge of their children in three ways: (1) parents produced longer, more informative referring expressions for animals that children generally learn later, (2) over and above this coarse-tuning, parents fine-tuned information to their individual children's knowledge of specific animals, and (3) when children did not know an animal that parents thought they did, parents subsequently produced longer referring expressions for that animal. Further, this tuning was associated with more successful communication: children were more likely to correctly select animals whose names parents thought they did not know if parents produced more informative referring expressions.

These data are consistent with prior evidence of coarse-tuning in child-directed speech, but importantly provide the first experimental evidence for fine-tuning at the lexical level. When communicating with their children, parents not only take into account the average difficulty of each animal word, but they also rely on (and update) their estimates of their individual child's knowledge of those animals. Coarse-tuning and fine-tuning may be distinct adaptations that happen independently at different timescales, but our data suggest an intriguing alternative possibility: parents' coarse-grained estimates of their children's language development may be built hierarchically. That is, parents may use estimates of their children's knowledge of specific lexical and syntactic items to form their general representations of their children's overall language development. Hierarchical representations are a powerful vehicle for maximizing both speed and generalizability of learning, and they may play the same role here, allowing parents to efficiently track and use their knowledge of their children's language development (Tenenbaum et al., 2011).

While parents' speech to children is unlikely to reflect an explicit goal to teach, it is nonetheless goal-oriented: parents want to communicate successfully (Bruner, 1983). The reference game was conducted in a laboratory, but similar communicative pressures structure the daily conversations between children and their parents (Tamis-LeMonda et al., 2017). When talking about animals that they thought their children did not know, parents used referring expressions rich with descriptors and comparisons, as in previous observational studies (Blewitt, 1983; Masur, 1997; Mervis & Mervis, 1982). These strategies facilitate communication–parents use what they think their children know (e.g., color words) in order to communicate about animals they think their children do not know. Because communication and learning are intertwined, these same strategies may work in the service of language acquisition (Yurovsky, 2018). While parents produced rich descriptions to help their children select unfamiliar animals, they almost always produced the canonical label as well. These referring expressions are thus an ideal opportunity to learn the relationship between the referent, its label, and its important identifying features. We did not independently measure children's knowledge of each animal, so we cannot determine whether they learned any new animals while playing the game. The relationship between referential strategies and ultimate learning is a promising direction for future work.

Parents fine-tune language to their children's knowledge to communicate successfully. In the service of proximal communicative goals, they may also provide children with input that ultimately accelerates learning. Focusing on the interactive and communicative nature of language captures a more complete picture of language development: while children bring powerful learning mechanisms to language acquisition, these mechanisms are supported by a linguistic environment designed for their success. Our evidence suggests that parents are highly sensitive to their toddler-age children's vocabulary knowledge, and tailor language for their specific children. How does this type of tailoring or adaptation change as children develop, and begin to actively participate in conversation with caregivers? In Chapter 2, I examine the role that parents play in scaffolding children's communication skills. What types of support to parents provide during referential communication, and how does it change over children's development?

Chapter 2. Parents scaffold the formation of conversational pacts with their children

Children don't just learn the sounds, words, and structures of their native language—they learn how to *use language to communicate* (E. V. Clark, 2009; Hockett & Hockett, 1960). While we often think of words as having fixed meanings, and referents as having fixed labels, this is an oversimplified view (Elman, 2009; Wojcik et al., 2022). The same referent can be called different labels in different contexts, and the same word may take on new *ad hoc* meanings with different partners. For example, the same entity can be referred to as a dog, a pug, an animal, or just Max (Brown, 1958).

Effectively navigating this vast space of referential choice requires conversational partners to be sensitive to *common ground*. We produce and interpret language for each other in light of our shared history, and expect others to do the same (Bruner, 1985; H. H. Clark, 1996). Conversations are thus an ongoing cooperative process; even simple referential expressions are established collaboratively. When the meaning of an utterance is initially unclear, or when an intended referent does not have a singular conventional label, interlocutors will engage in negotiation, arriving at a shared *conversational pact* about how to think and talk about the intended referent (H. H. Clark & Wilkes-Gibbs, 1986; Haber et al., 2019; Hawkins et al., 2020).

These pacts are temporary agreements about referent names, and adults fluently form, revise, and track them over time with different partners. For instance, Brennan and Clark (1996) showed that adults will use the contextually appropriate level of specificity for a referent (e.g. "loafer" rather than "shoe" when another shoe is present), carry it forward into a new context with the same conversational partner even when the initial competitor is removed, then revert to "shoe" when a new partner is introduced. Conversational pacts are thus *partner-specific*, suggesting that pacts are formed for the purpose of in-the-moment communicative success (Brown-Schmidt, 2009; Brown-Schmidt et al., 2015).

While adults readily form conversational pacts with one another, children struggle substantially both to establish and to deploy conversational pacts with peers. In one set of studies, Krauss and Glucksberg (1977) asked pairs of age-matched children to play a director-matcher game that required them to refer to novel objects (see also Glucksberg et al., 1966; Glucksberg & Krauss, 1967; Krauss & Glucksberg, 1969). Across repeated interactions with the same novel objects, 5- and 6-year-old children showed very little evidence of pact formation: The matcher's accuracy in selecting the correct referent remained low, and the director did not reduce the length of their referential expressions as adults do. Older children showed some improvement on both measures, but even 10- and 11-year-olds were less successful than a comparison group of adults. These results are perhaps surprising, given that children engage in conversation much earlier, and are able to track at least some aspects of partner-specific common ground (Akhtar et al., 1996; Yoon et al., 2021).

Why, then, do children struggle to collaboratively construct conversational pacts with one another? In order to succeed in establishing reference, an utterance must solve several basic problems. First, there is an absolute problem of lexical meaning: the utterance must be recognized as a good enough literal description of the target object for the listener to identify it as applicable. Second, there is a relative problem of informativity: it must be recognized as a poorer description for other potential referents (distractors) in the context than it is for the referential target. Third, establishing reference requires work beyond the single-utterance level; mutual understanding must be monitored and maintained at the discourse level. That is, it is not enough to produce applicable and informative descriptions, interlocutors must be able to signal whether or not they actually understand one another in order to resolve misunderstandings. Thus, developmental problems in pact formation could stem from three potential sources of difficulty:

Utterance-level comprehension: One possibility is that children are unreliable listeners who are unable to identify the intended target even given a sufficiently informative utterance. In order to identify the target of an utterance, children need to be able to determine whether the information in it is a good description of each potential referent. Children may thus struggle to identify the relevant and informative components of a complex utterance like "the shoe with the brown laces" and take interpretive shortcuts. For example, they may think that any referent that matches a part of the utterance like "shoe" is good enough and guess (Speer, 1984).

Utterance-level production: Alternatively, the primary source of difficulty may lie on the production side. Children may generate referential expressions that are either overly idiosyncratic or insufficiently informative for their conversational partner to correctly identify the intended referent (Beal, 1988; Beal & Flavell, 1983). Idiosyncracies may arise for novel objects, such as those studied by Krauss and Glucksberg (1977), due to difficulty identifying which features to refer to without a prior history of hearing adults talk about the target object. For familiar objects, conventional labels may provide a good default, but conventional labels can be insufficient in more challenging contexts, such as when a shoe should be called a "loafer" because of the presence of other shoes (Brennan & Clark, 1996; Goldberg & Ferreira, 2022; Koranda et al., 2022). Insufficient informativity is also consistent with evidence suggesting that young children may struggle to determine which referential expressions are more or less informative in context (Beal & Flavell, 1982), even for idiosyncratic expressions they themselves produced at an earlier time (Asher & Oden, 1976; Robinson & Robinson, 1977).

Discourse-level interaction: Finally, the difficulty may lie beyond the level of single utterances, implicating the interactive discourse processes that link together production and comprehension. For example, children may experience no more or less difficulty in production or comprehension than adults, but struggle to notice and actively signal to their partners that they have not understood each other. Adults readily use back-channels and other vocal affirmations to signal that they have understood a potentially ambiguous utterance, and will ask clarifying questions when they feel they have been unable to understand (H. H. Clark & Wilkes-Gibbs, 1986; Schegloff, 2007). However, young children may have trouble explicitly initiating these speech acts when necessary, and errors may persist preventing pacts from getting off the ground (Anderson et al., 1994; Beal, 1987).

In a series of three studies, we investigate the contributions of these three sources to previously observed difficulties in referential pact formation among US children. Our investigation begins with the observation that, even if young children are known to struggle to form pacts with one another, they nevertheless spend many hours in conversation with parents and other caregivers. It is possible that adults, themselves fluent conversationalists, are doing something to support the formation of conversational pacts when talking with their children that children are not able to do with one another; thus, successes or failures in the parent-child setting may shed light on the sources of previously reported failures in the child-child setting.

In Experiment 3.1, we adapted the classic tangram director-matcher paradigm developed by H. H. Clark and Wilkes-Gibbs (1986) to examine conversational pact formation among parent-child dyads. We found that even 4-year-old children and their parents can successfully coordinate on pacts, but for these youngest children, parents both provide significant interactional support and are the original source of most of the labels that go on to constitute successful pacts. The rest of the paper aims to disentangle the contributions of children's comprehension and production skills. In Experiment 3.2, we isolated the role of comprehension processes by asking naive adults and children to identify the target reference given the utterances originally produced in our parent-child interactions. In Experiment 3.3, we isolated the role of production by asking children and adults to describe familiar and novel objects in different contexts. We found that comprehension is relatively constant across development, but younger children produce both more lexically idiosyncratic and less context-sensitive labels. These findings suggest that impaired production may prevent pacts from taking root in pairs of young children, but when adults are able to provide enough interactional scaffolding, even very young children are able to adapt and establish common ground.

2.1. Experiment 2.1: Forming pacts in an interactive communication game

In our first experiment, we explore whether children may be able to coordinate on conversational pacts with adult partners, even as numerous studies have revealed difficulties with peers. Adults may provide children with a model of informative referential expressions, and actively seek clarification when children's utterances are ambiguous (E. V. Clark, 2018; Nikolaus et al., 2022). It is clear from existing studies that children are receptive to *explicit* coaching from experimenters. For example, Deutsch and Pechmann (1982) showed that repeated clarification questions eventually elicited unambiguous descriptions from 6- and 9-year-olds and, to a lesser extent, from 3-year-olds. A more recent training study by Matthews et al. (2007) replicated this finding in 2- to 4-year-olds, and further showed transfer to a different communication task. But it is unclear whether, or how, adults spontaneously provide such feedback in the course of natural adult-child conversations, and whether such processes actually enable the formation of pacts (rather than generically increasing informativeness).

One possibility is that caregivers naturally adopt some of the interactive strategies developed by experimenters in these training studies: pointing out when references are insufficiently informative, asking for clarification about targeted features, and helping children to re-conceptualize referential targets within the expressiveness of their developing vocabulary. In this study, we asked parents and their 4-, 6-, and 8-year-old children, as well as control pairs of adults, to play an adapted version of Krauss and Glucksberg's (1969) director-matcher game. We examined (1) whether parents' contributions enable conversational pacts to form at all, (2) what strategies parents use to scaffold pact formation if so, and (3) whether the nature of these pacts changes over development as children become more adult-like users of language.

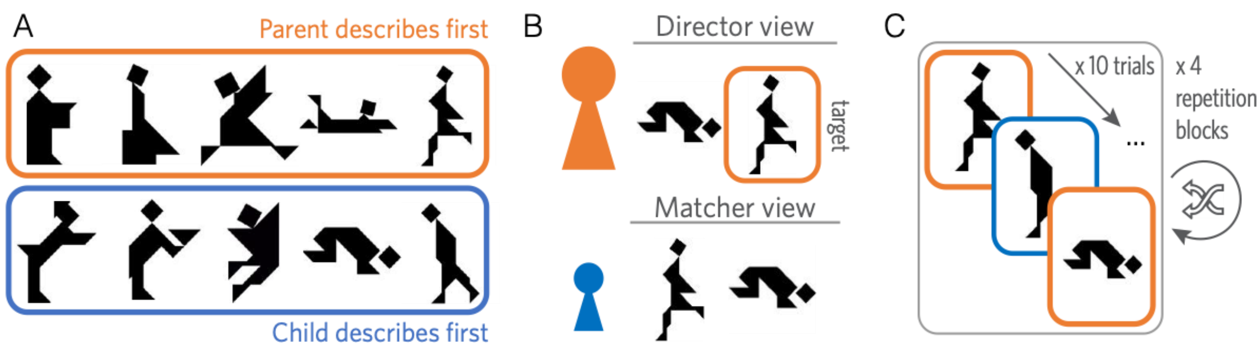

Figure 2.1 *(A) We divided a set of ambiguous tangram figures into two sets of referential targets. (B) A parent (orange) and child (blue) played a reference game where these figures were presented as stimuli. On each trial, the director (here, the parent) was asked to refer to the target figure (privately highlighted in a box) so that the matcher (here, the child) could guess it. (C) Each tangram appeared as the target once per black, and each dyad played the game for four blocks, alternating roles on each trial.*

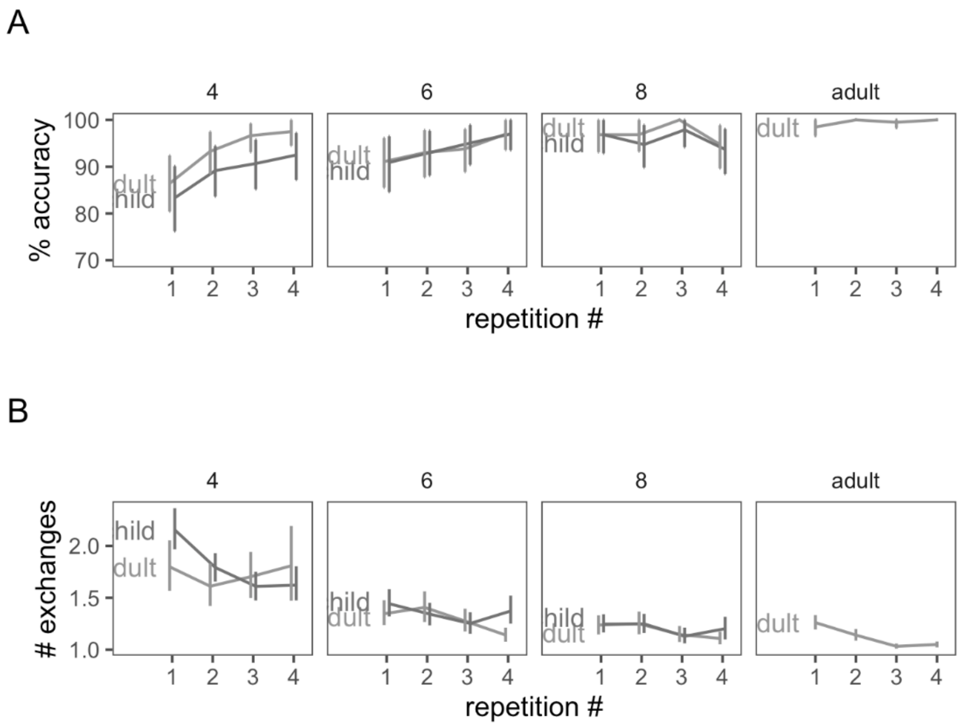

Figure 2.2 *(A) Accuracy and (B) number of dialogue exchanges per trial, broken out by whether the child (dark grey) or the adult (light grey) was the director. Error bars are 95% CIs. The drop in accuracy for 8-year-olds is likely an artifact of a small set of children losing interest in the task.*

2.1.1. Methods

Participants

Children (ages 4, 6, and 8) and their parents were recruited from a database of families in the local community to achieve a planned sample of 60 parent-child pairs (20 per age group). A total of 75 children and their parents participated. Data from 12 pairs were dropped due to failure to complete the study, leaving a final sample of 63 pairs. For comparison, adult participants were also recruited from a Psychology Department subject pool to achieve a planned control sample of 20 adult-adult pairs.

Stimuli

Twelve solid black images of tangrams were normed on Amazon Mechanical Turk for pairwise similarity. Each of the 60 participants made 22 pairwise similarity judgments on a scale from 1-100. Based on these similarity ratings (M = 42.3, SD = 26.1), the ten most dissimilar tangrams were selected for use as stimuli (**Figure 2.1**).

Design and Procedure

Pairs of participants were brought into the lab to play a cooperative director-matcher game. They were seated in front of iPads at opposite ends of a table, with a divider preventing them from seeing each other's screens. Participants were told that they would take turns playing *director* and *matcher* roles. The director's task was to describe the *target* image, privately indicated by a blue border, and the matcher's task was to select one of the images on their screen based on the director's description (**Figure 2.1**). Before beginning the experiment, participants played six practice trials with images of common fruits and vegetables.

The experiment consisted of four repetition blocks of ten trials each (**Figure 2.1**). Each tangram was the target once per block. We constructed the trial sequence to ensure that participants both alternated roles from trial to trial and alternated roles for each *target* from block to block. For each participant pair, we randomly divided the tangrams into two sets of five: the adult was assigned one set to describe on the first block, and the child was assigned the other set. These sets were interleaved on the first block, such that players alternated roles. On each subsequent block, these sets were swapped such that each tangram was described by each participant exactly twice over the course of the experiment.

On each trial, the target tangram appeared with a foil selected from the set of nine other tangrams. Targets appeared with different foils on different repetition blocks. To ensure that the game would not be too difficult for young children, tangrams most similar to the target (based on similarity norms) did not appear as foils. To discourage participants from using spatial language (e.g. “left side”), the target and foil were shown in randomized order across the two iPads. When the matcher selected an image, it became colorful and a pleasant sound played. Importantly, neither the matcher nor director received feedback about accuracy: the same sound played whether the selection was correct or not.

Pre-processing

Sessions were videotaped and subsequently transcribed using Datavyu (2014), an open-source coding program. Each video was transcribed and checked by two different coders. Utterances were manually coded as part of a trial or unrelated to the game (e.g., “sit down please”), and unrelated utterances were removed before analysis.

2.1.2. Results and Discussion

We characterized developmental differences using three measures of communicative behavior. First, we examined accuracy to evaluate whether children were able to succeed at the reference game in collaboration with their parents. Second, we examined conversational turn-taking behavior to evaluate how interactive dialogue may contribute to success. Third, we examined the number of words produced by each partner on each turn to evaluate the efficiency of pacts.

Performance

We began by analyzing task performance across age groups. Because pairs of adults were consistently at ceiling throughout the task, we focused on the performance of parent-child pairs. We constructed a mixed-effects logistic regression predicting whether the matcher successfully chose the correct referent on each trial. The model included fixed effects whether the parent or child was the director, (numeric) age, and repetition block. It also included random intercepts for each tangram and pair of participants, and random effects of repetition block and director for each pair of participants.

Initial accuracy was well above chance for all age groups, indicating that even young children can succeed in this referential task with their parents. We also found a significant main effect of age (β = 0.34, t = 2.99; $p = .003$): Pairs with younger children performed significantly worse than pairs with older children. Critically, however, accuracy improved significantly over the four repetition blocks for all groups (β = 0.49, t = 3.48, p < .001; **Figure 2.2**). Such improvement for 4-year-olds contrasts with previous results showing no improvement in accuracy with pairs of kindergarteners (age 5, Krauss & Glucksberg, 1969). Intruigingly, accuracy was lower when children were the directors (β = -0.38, t = -2.18; $p = .030$), suggesting a potential asymmetry in performance across roles.

Interactive dialogue

If the ability of children of different ages to successfully establish reference depends on interactive scaffolding provided by their parents, we would expect additional dialogue exchanges for younger children. We quantified dialogue exchanges by counting the total number of distinct turns of continuous speech on each trial and constructed a mixed-effects model predicting the (continuous) number of exchanges with the same effect structure reported in the previous section. Consistent with previous work, and replicated in our adult control condition, we found a significant main effect of repetition: Fewer dialogue turns were required on later trials [β = -0.07, t = -3.13; p .003; H. H. Clark and Wilkes-Gibbs (1986)]. In line with our predictions, we also found a significant main effect of age (β = -0.14, t = -3.47; p < .001). Pairs with 4-year-old children took roughly one additional turn at each point in the experiment than pairs with older children, who more closely resembled pairs of adults (**Figure 2.2**).

The increased levels of interactivity between parents and younger children provides an interesting contrast with previous studies showing decreased interactivity between peer dyads of younger children (Anderson et al., 1994). These lengthier exchanges parent-child exchanges may reflect efforts by parents to provide and elicit additional clarification or confirmation, or may simply reflect attentional difficulties. As a rough estimate of the content of parents' responses when their child was the director, we counted the number of question marks in the transcript[1]. We found 44.37% of responses to 4-year-olds contained question marks, compared to lower rates of 27.57% and 31.01% in 6- and 8-year-olds, respectively ($\chi^2(2) = 22.68$, $p < .001$).

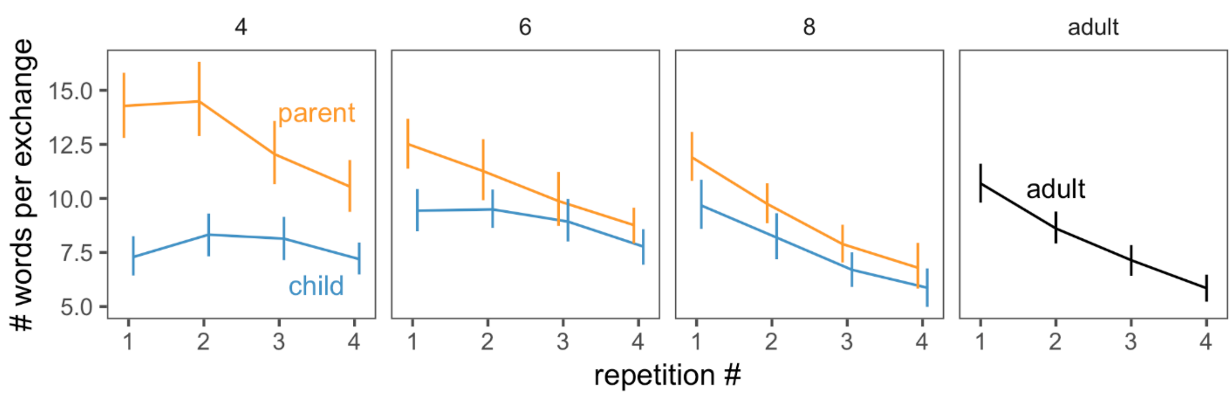

***Figure 2.3** Total number of words in referential expressions produced by children and parents over the course of interaction.*

Reduction in length of referential expression

A key signature of successful communication among adults is an increase in efficiency over repeated reference (H. H. Clark & Wilkes-Gibbs, 1986). As pairs form conceptual pacts, they

[1] Because questions were not always annotated with a question mark, this represents a lower bound on the true estimate of questions.

are able to communicate the same meaning using fewer words. Our control sample of adults replicated this classic effect ($\beta =$ -0.23, $t =$ -6.21, < .001). Here, we asked whether parents and children of different ages spontaneously reduce their referential expressions in the same way. We define referential expressions as all utterances produced by the director on a given trial, up until a selection is made by the matcher. Because the total number of words produced on a trial is correlated with the number of dialogue exchanges examined above ($r =$ 0.60), we constructed a normalized efficiency measure that controls for additional turns. Specifically, we defined efficiency as the total number of words produced by the *director* divided by the number of dialogue exchanges on that trial.

Because participants in each pair alternated roles, each participant served as the director twice for each tangram. This structure allowed us to examine how each participant changed their language when they were the director (**Figure 2.3**). Using a mixed-effects model, we predicted the (log) number of words per exchange on each trial, including fixed effects of age, repetition block, and speaker identity (parent vs. child) as well as all of their interactions. We also included random intercepts at the tangram-level and maximal random structure at the dyad level (i.e. intercept, slopes for repetition block and speaker identity, and their interaction; Barr et al. (2013)). All variables were centered to allow interpretation of lower-order terms as effects at the average level of the other terms. We found significant main effects of repetition block ($\beta =$ -0.13, $t =$ -6.21, p < 0.001), speaker identity ($\beta =$ 0.15, $t =$ 7.09, p < 0.001), and age ($\beta =$ -0.05, $t =$ -2.06, p = 0.043). All else being equal, directors used fewer words over subsequent repetitions, children used fewer words than their parents, and pairs with older children used fewer words than pairs with younger children. However, these main effects were clarified by several interactions of interest.

First, while parents on average used more words as director than their children did, we found a significant interaction with the child's age ($\beta = -0.04$, $t = -3.15$, $p = 0.003$). This gap between parent and child utterance length was largest at age 4 but nearly disappeared by age 8. Second, we found that parents reduced their utterance length over time more strongly than children did, holding age group constant ($\beta = -0.03$, $t = -3.74$, $p < 0.001$). Third, having older children in a pair supported stronger reduction overall, ($\beta = -0.04$, $t = -3.80$, $p < 0.001$). An intriguing final question is whether the *rate of reduction* changes over the course of development: are 4-year-olds able to become more efficient as common ground is built in the same way 8-year-olds do? Because parents of different age groups display roughly similar slopes of reduction, this question is addressed by a 3-way interaction. We found that this interaction was not significant ($\beta = 0.01$, $t = 1.81$, $p = 0.075$), although our sample was likely underpowered to detect this higher-order interaction.

Where do pacts come from?

Our results so far demonstrate that children are able to converge on increasingly accurate and efficient pacts with their parents. What might allow children to coordinate with their parents but not with their peers (Krauss & Glucksberg, 1977)? A classical explanation is that children are rigid and lack the ability to *adapt* to their partner: they have a strong preference for a particular idiosyncratic description and are not sensitive to the possibility that their partner may not understand it (e.g., "this one looks like mommy's dress''). Under this hypothesis, children fail with other children because they each stubbornly continue to use mutually incomprehensible expressions, and only succeed with their parent as a result of the parent's flexibility.

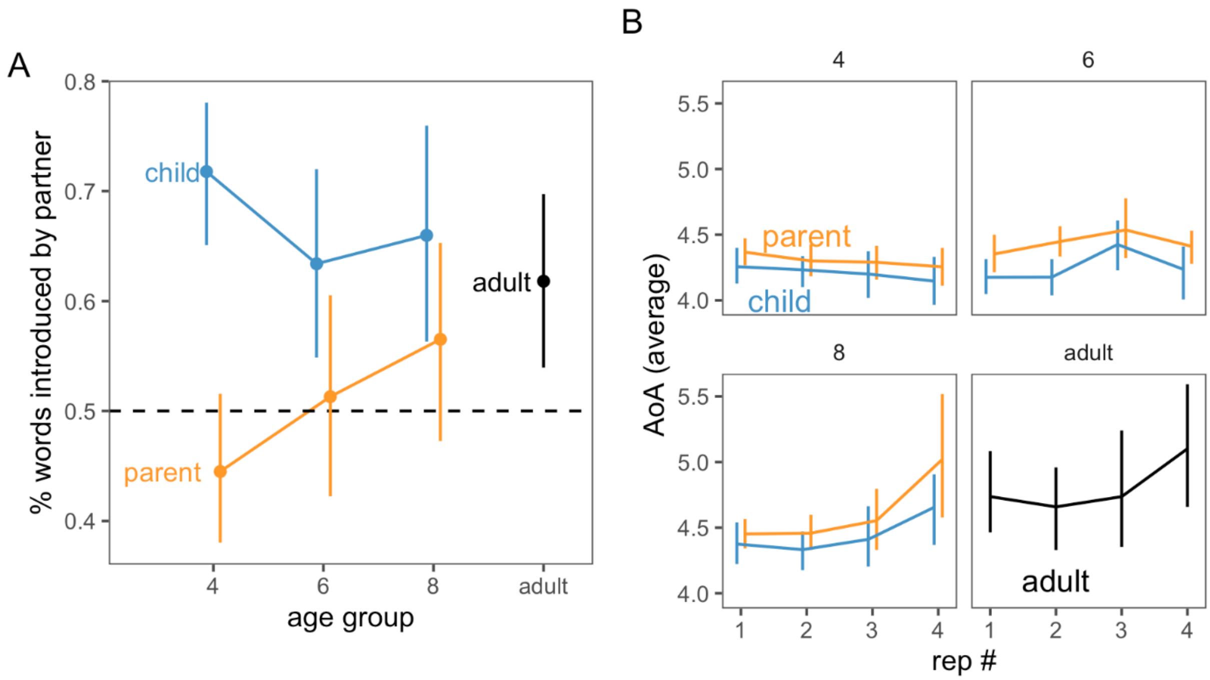

Figure 2.4 *(A) Probability of words used on final round first occurring with child or parent. (B) Complexity of language used by different age groups estimated from words' average age of acquisition. Error bars are 95% CI.*

Another possibility is that young children may be able to adapt successfully but are simply unable to generate good enough initial candidate labels to get the process off the ground. In this case, pairs of children may fail because neither partner can generate good enough labels to start the pact-formation process, while children and parents succeed because parents seed the first good candidate label. Each of these accounts make different predictions about who is adapting to who: do pacts originate with children, or with adults? We distinguish these accounts by quantitatively analyzing the natural-language transcripts.

For each word in the final description of a tangram, we checked whether it had appeared in an earlier referential expression for that tangram. We noted the first trial where it appeared, and who was director when it was produced[2]. The proportion of words originating with the child and parent is shown in **Figure 2.4**. We observed an asymmetry: the words used by children on the final repetition were more likely to have originated with their parents than the words used by parents were to originate with their children. In addition, this gap appeared to close with older groups, with parents more likely to adopt words introduced by older children.

We tested this hypothesized interaction using a mixed-effects logistic model predicting whether each word appearing on the final repetition for each tangram was introduced by the current director or by their partner. We included fixed effects of age group and the speaker identity (parent or child), as well as random intercepts for each pair of participants and each tangram. We found a significant main effect of speaker identity, with the words used by children more likely to originate with their partner than the words used by parents, ($\beta = 0.37$, $t = 4.76$, < .001). Additionally, we found a weak but significant interaction between speaker and age, indicating that this asymmetry was smaller for older children ($\beta =$ -0.10, $t =$ -2.19, $p =$ .028). Thus, parents—especially parents of younger children—appear to be the source of the words that persist in successful conceptual pacts.

How do pacts differ across ages?

Given that the labels used by younger children tend to originate with their parents, it is tempting to conclude that adaptation for these pairs is entirely one-sided. However, the resulting pacts may have nonetheless been shaped by children. In particular, we suggest that even though young children may not be able to *supply* good initial labels, they may serve as "simplifiers" or

[2] To match different forms of the same word (e.g. "jumping" vs. "jumped") we first lemmatized each word. We also filtered out stop words ("the", "with"), as well as common words that were not part of the pacts ("person", "box"), and excluded words that appeared for the first time on the final repetition of each target.

“filters” that constrain the complexity of the resulting pact, since it must pass through their comprehension and production system (Hudson Kam & Newport, 2005). We tested this hypothesis by asking about the linguistic complexity of the pacts developed in each age group. To estimate the complexity of individual words, we used self-reported age of acquisition (AoA) estimates (Kuperman et al., 2012). Our assumption, based on Kuperman et al.’s work, is that more complex words are harder to acquire, and thus have later ages of acquisition (AoA). To estimate the complexity of referential expressions as a whole, we simply averaged the AoAs of all the words they contained[3]. In all age groups, parents produced more linguistically complex referential expressions than children, but as children get older, both their expressions and their parents’ become more linguistically complex over repetitions of the same tangram (like adults) (**Figure 2.4**).

We confirmed this observation with a mixed-effects model including all possible interactions of age, person (child vs. parent), and repetition number, and maximal random effect structure as before. We found significant main effects of age ($\beta = 0.06$, $t = 2.49$, $p = 0.016$), and person ($\beta = 0.15$, $t = 4.59$, $p < 0.001$), but not repetition ($\beta = 0.03$, $t = 1.55$, $p = 0.127$). All else being equal, pairs with older children produced more complex utterances, and parents produced more complex utterances than children. Further, we found a significant interaction between age and repetition number, suggesting that later repetitions became more complex for dyads with older children ($\beta = 0.03$, $t = 2.76$, $p = 0.008$). This interaction may be driven by reduction processes, as more distinctive words tend to persist as common words are dropped (Hawkins et al., 2020). No other effects were significant, indicating that parents’ utterances remained a constant level of

[3] We first removed any words in a list of 174 stop words like “at” and “me” from the package (Feinerer et al., 2008)

complexity above children's even as children's utterances became more complex. These results suggest that pact formation depends not only on parental input, but on children as well—pacts may originate from parents, but children control their complexity.

2.2. Experiment 2.2: Comprehension

Our first experiment found that young children are able to coordinate on referential pacts with their parents. Even four-year-old children readily adopted the labels introduced by their parents and interactively refined their descriptions in response to spontaneous parent-initiated repair. Overall, four- to eight-year-old children and their parents exhibited patterns similar to adult pairs in terms of reduction in length of referring expressions and turns. Yet these findings alone do not resolve the puzzle of why younger children struggle so much with their peers. What are the underlying reasons for children's difficulties?

One possibility is that the difficulty lies primarily in the process of *comprehension*. That is, children might be poor listeners. They may be unable to understand or accommodate their partner's description when it does not align with their own way of conceptualizing a tangram, preventing uptake of a pact. In this case, parent-child dyads' communicative success in our reference game may be due to parents' ability to align to their children's points of view, such that children can accurately select the intended referent based on their parent's descriptions. Another possibility is that the root of the difficulty primarily lies in *production*. In other words, children may have difficulty *generating* sufficiently descriptive referring expressions on their own, given their limited vocabulary and other processing constraints, but are able to recognize a good description when they hear it, and adopt that description going forward. Further, children may have difficulty recognizing that their expressions are not informative enough for their conversational partner.

In Experiment 2, we explicitly tested the first of these two hypotheses. We provided naive children and adults with the descriptions produced by participants in Experiment 1 and evaluated how well they were able to interpret them. If comprehension is the root of the difficulty forming pacts, we might expect that naive children as comprehenders would be equally unable to interpret all referring expressions (regardless of whether they were originally produced by adults or children) while naive adults as comprehenders would have no difficulty. Conversely, if naive children are able to understand the referring expressions produced by parents (though not necessarily those by children), then we may expect the source of the difficulty to lie elsewhere.

2.2.1. Methods

Participants

We recruited 355 adults from Amazon Mechanical Turk. Data from 24 participants were dropped due to failure to complete the study, leaving a final sample of 331 adults. Additionally, we planned to recruit a sample of 200 children (ages 4 to 8) from a school and a museum in the local community. However, because of the COVID-19 pandemic, we were forced to terminate data collection early with only 78 children (ages 4-8). Data on race and ethnicity were not collected, but we expect our participants to reflect the racial/ethnic makeup of the Chicagoland area.

Stimuli

To conceal the identity of the original speaker, the first author and four research assistants produced new audio recordings by reading the Experiment 1 game transcripts in a uniform vocal style. All recordings were done by female native English speakers. To examine the extent to which pacts became more opaque over time, we used utterances from the first and fourth (final) round of the reference game. We removed disfluencies and isolated the speaker's original referring expression on each trial (i.e., we excluded additional information provided in response to the listener's questions or prompting). This process produced unique audio stimuli, two for each item of each game from Experiment 1. These stimuli were broken into 118 unique 'item sets' containing 10 recordings each, such that each tangram appeared as the target once in each set. On each trial, the target tangram appeared alongside its foil from the corresponding trial in Experiment 1. The stimuli sets were designed such that all participants encountered 5 utterances generated by a parent and 5 by a child, and 5 utterances from the first round and 5 from the final round.

Procedure

Participants were placed in the role of the listener and presented a sequence of 10 audio recordings — a single referring expression for each target tangram, in a randomized order. Participants were instructed to click the intended referent based on the audio they hear. On each trial, two tangram images were displayed side by side (left and right order randomized). At the beginning of each trial, the audio recording played once. To reduce possible learning effects, participants did not receive any feedback after their response. Before participants began the experiment, we ensured their audio was working. We did not allow participants to proceed past the consent page without clicking a 'play' button that asked them to "type the number 86 into the box." Additionally, to detect participants who were not following instructions, we included an attention check which simply asked people to click 'the one on the left'. Child participants provided verbal consent prior to the start of the experiment, and parents provided written consent. Children's version of the experiment did not include the initial 'play' button and attention check. While children completed the task independently, an experimenter was nearby to ensure that the audio was working. We measured accuracy and response time, measured as the time interval between the completion of the audio recording and the response. Participants were not allowed to respond prior to the completion of the audio recording.

Design

We used a 2×2 factorial design manipulating the age of the producer (parent vs. child) and comprehender (adult vs. child). The age of the producer was a within-subjects manipulation while the age of the comprehender was an across-subjects manipulate: each comprehender was exposed to utterances originally produced by both adults and children. We predicted that if children struggle to form conceptual pacts primarily due to *comprehension* difficulties, then adult comprehenders should be highly accurate across the board (regardless of whether the referring expression was originally produced by an adult or child), while child comprehenders should uniformly struggle.

2.2.2. Results

Children and adults comprehend descriptions at similar levels

To test whether age effects in reference game performance may be attributed to comprehension difficulties, we compared the extent to which a naive group of children and adults were able to determine the referent of the same expressions. We focus primarily on accuracy, as response times are not directly comparable between the web interface (for adults) and lab interface (for children; see Supplemental). We constructed a mixed-effects logistic regression model predicting trial-by-trial accuracy, including a fixed effect of comprehender group (child vs. adult) and random intercepts and slopes for each source game[4]. We found no significant difference in overall accuracy across groups, (adult accuracy $=$ 0.87, 95% CI $=$ [0.86, 0.88], child accuracy $=$ 0.85, 95% CI $=$ [0.83, 0.88], $t = 1.49$, $p = 0.14$, see **Figure 2.5**). Consistent with this finding,

[4] Because we terminated data collection early, we do not have child comprehension data for all audio clips, preventing us from including random slopes at this finer level of granularity. All behavior within a game is causally intertwined, so we believe this coarser level of grouping is most natural.

there was no support in a nested comparison for a model with an effect of comprehender group over an intercept-only model, $\chi^2(3) = 2.28$, $p = 0.52$.[5]

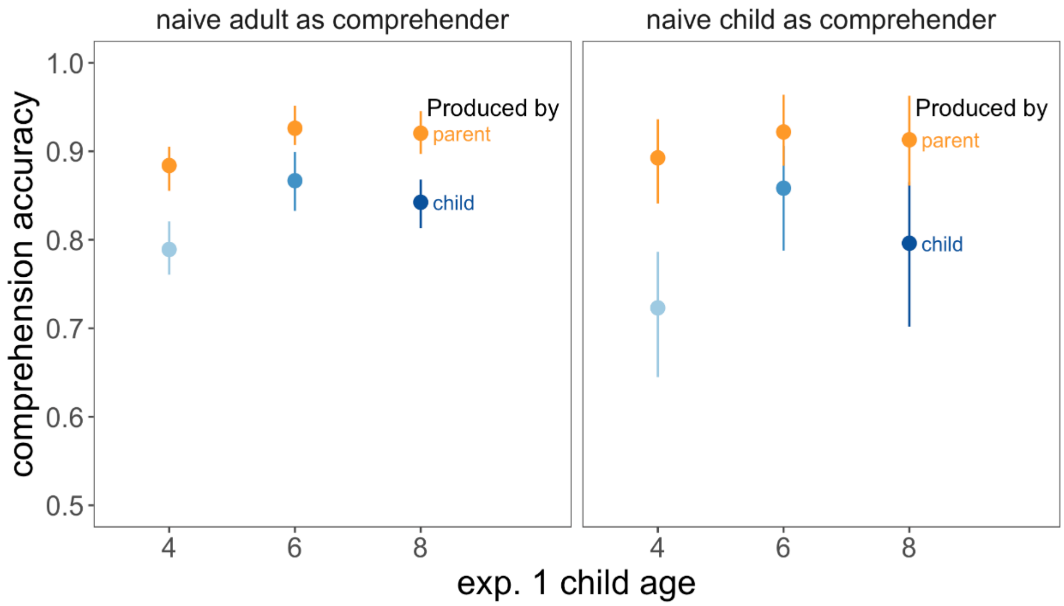

Figure 2.5 *Results for Experiment 2. Both naive adults (left) and children (right) were able to more accurately interpret referring expressions originally produced by parents (orange) than children (blue) in Experiment 1. No overall difference was observed across the two comprehender groups. Error bars are 95% confidence intervals.*

Which descriptions are easier to comprehend?

Next, we examine which features of the descriptions may affect downstream comprehension. We consider two specific hypotheses. First, we predicted that descriptions produced by adults will be easier for *everyone* to comprehend, overall, than descriptions produced by children in the same dyad. Second, we predicted that descriptions produced earlier in the game will be easier to comprehend than descriptions produced later in the game, consistent with prior

[5] Because null hypothesis significance testing is unable to provide positive evidence for the null hypothesis, we also ran a Bayesian regression using the brm package. We used the Savage-Dickey method and found a Bayes Factor of 9.5, indicating moderate support in favor of the null hypothesis $b = 0$. We used a weakly informative student-t prior on the coefficient with degrees of freedom $\nu = 5$ and scale $s = 2.5$, following Gelman et al. (2008), but this result was robust to other choices.

findings that pacts are not always transparent to eavesdroppers (Schober & Clark, 1989). Because we found no overall differences in comprehension group in the previous section, we collapse across both adult and child comprehenders for the following analyses.

To test these predictions, we fit a logistic mixed effects model predicting accuracy as a function of the person who produced the description (parent vs. child), the age of child in the pair (ages 4, 6, 8), and the phase of the game where it was produced (early vs late). We also include random intercepts for each target and participant ID; a more complex model adding interactions was not justified, $\chi^2(4) = 4.35, p = 0.36$. How does age of speaker influence comprehension of descriptions? We first find that parents' descriptions are more accurately comprehended across the board $b = -0.40, z = -8.16, p < .001$. This is in line with our predictions, suggesting that parents in Experiment 3.1 provided more informative descriptions of the tangrams than did children. Additionally, we find that older children (and parents of older children) provided more accurately comprehended descriptions than younger children, $b = 0.19, z = 3.82, p < .001$ (**Figure 2.5**). Concerning our second hypothesis, however, we found no significant effect of experiment phase, $b = 0.00, z = 0.04, p = 0.96$. In other words, final-round descriptions of tangrams do not appear to contain meaningfully less information for guessing the target.

2.2.3. Discussion

In Experiment 3.2, we asked whether children's struggles to form conceptual pacts with others are due to comprehension-side difficulties. Surprisingly, however, we found no difference between naive adults and children's ability to comprehend referential expressions generated by participants in Experiment 3.1. Although children may not necessarily be worse at comprehending utterances, our data did reveal that referential expressions originally *generated* by younger children were less accurately comprehended by naive adults and children alike. That is, both adults and children were less likely to find the intended referent after hearing a description originally produced by a 4-year-old child.

If children's referential expressions are difficult for naive listeners to comprehend, and late-stage utterances were no more or less comprehensible that early-stage ones, then what contributed to communicative success among parent-child dyads in Experiment 3.1? One likely explanation for parents' higher accuracy in the interactive game is that they were able to actively seek clarification from children in order to select the correct referent when children's utterances were ambiguous. In our comprehension task, audio recordings comprised only of initial expressions for each round — any further exchanges and clarification were removed. This mechanism is also consistent with our observation that parents are more likely to ask questions when their young children were directors. From these findings, we might infer that 4-year-old children's success in our interactive task (Experiment 1) relied on some degree to their parents' ability to overcome any idiosyncratic or ambiguous descriptions generated by young children. Taken together, then, our findings suggest that children's difficulty forming pacts may lie outside their comprehension processes and point toward possible developmental changes in children's ability to produce sufficiently informative descriptions.

2.3. Experiment 2.3: Production

In Experiment 2.2, we found that naive children and adults were able to comprehend referential expressions equally well. Contrary to a classical view that young children are rigidly fixating on a single way of conceptualizing the tangram, even young children seemed relatively flexible in their ability to accommodate different descriptions. That is, to the extent that children contributed to poorer group performance in Experiment 2.1, their contribution is not well-explained by the *comprehension hypothesis*. At the same time, however, we found some evidence consistent with the *production hypothesis*: messages originally produced by children were somewhat less comprehensible for listeners of any age (as indexed by slower reaction times and lower accuracy).

While this source effect is intriguing, it is difficult to disentangle the original messages from the interactive parent-child context in which they were produced. For example, it is possible that children were receiving interactive scaffolding that prompted them toward more comprehensible expressions. Or, conversely, it is possible that children were relying on their parents to take on more of the division of labor of interpretation (Hawkins et al., 2021) and were thus producing expressions that are *less* comprehensible to naive audience than they were actually capable of. In Experiment 3, we remove the interactive context to more directly assess children's ability to produce referential expressions for novel tangram objects in different contexts.

We used a 2x2 design aiming to tease apart two related explanations for poor production performance. First, to assess the extent to which production difficulties stem from *pragmatic reasoning* (i.e., the ability to recognize that an accessible label is not sufficiently informative in context), we manipulate whether the distractor in context is more or less similar to the target. Second, to assess the extent to which production difficulties stem from impoverished *lexical priors* (i.e., the ability to access candidate labels for a given referent), we manipulate whether the target objects are familiar photographs or novel tangram shapes (see Horton & Gerrig, 2002). To the extent that children fail to produce context-sensitive utterances for familiar objects with accessible labels, we may expect that pragmatic reasoning is the primary bottleneck on performance. To the extent that there is greater overall variance in the utterances produced by children, we may expect that lexical priors play a larger role.

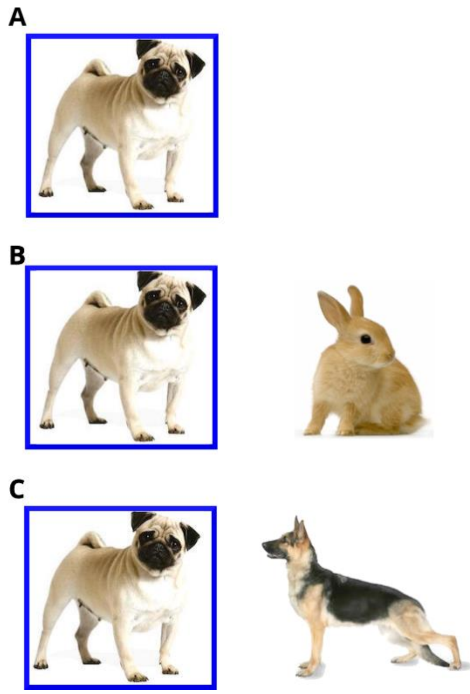

Figure 2.6 *Objects in Experiment 3 could appear in isolation (A), in a far context (B), or in a close context (C).*

2.3.1. Methods

Participants

We recruited 100 adult participants from Amazon Mechanical Turk. All participants gave informed consent prior to the start of the study and were compensated 60¢. We also recruited 60 children aged 4 to 8 years old ($M = 6.35$) to participate in the study. Data on race and ethnicity were not collected, but participants were recruited from a database of families that generally reflect the racial/ethnic makeup of the Chicagoland area. The study was conducted online over Zoom. Parents provided informed written and verbal consent, and children provided verbal consent. Families were compensated $5 for their participation.

Stimuli and Design

We used 37 pictures of familiar objects and 37 pictures of tangrams. Familiar objects were drawn from an image set used by Degen et al. (2020). We used eight different basic-level categories that we expected to be familiar to children: bears, birds, cars, candy, dogs, fish, shirts, and tables. Tangrams were drawn from a publicly royalty free set available on https://www.1001freedownloads.com/.

For both familiar objects and novel tangrams, trials were divided into three types: isolation trials in which a single target object was presented (**Figure 2.6A**), far trials in which a target was paired with a competitor of the same type but with low semantic overlap (**Figure 2.6B**), and close trials in which the target was paired with a competitor with high semantic overlap (**Figure 2.6C**). For familiar objects, close trials involved two objects in the same basic-level semantic category (e.g., Pug and German Shepherd), and far trials involved two objects in different basic-level categories (e.g., Pug and Rabbit). For novel tangrams, close and far competitors were determined on the basis of their degree of labeling overlap with target in a norming experiment. In this norming experiment, we asked adult participants to produce labels for a number of tangrams in isolation. We used naming agreement–the proportion of responses that were the same for two tangrams–to estimate their semantic similarity (see Zettersten & Lupyan, 2020), and then constructed close and far contexts based on high or low similarity, respectively.

Adults provided labels for a total of eight familiar objects and eight novel tangrams. Each object appeared three times: once in isolation, once in a close context, and once in a far context. All trials of a single type were presented in a single block (e.g., six trials in a row of 'familiar' objects in 'close' contexts), and the six blocks were presented in random order across participants. Children provided labels for the same eight familiar and novel objects. To maintain children's attention for the duration of the experiment, they gave responses only in the close and far conditions and not in the isolation condition. Children's productions in all conditions were typed into a text box in real time by an experimenter during their participation.

Children also participated in an additional task designed to measure individual differences understanding ambiguity in *comprehension* (Beal & Belgrad, 1990; Markman, 1977), as failures to judge the informativity of others' expressions may be implicated in failures to produce appropriately informative utterances of one's own in the close condition (e.g., saying "bear" when both the target and distractor are bears). In this condition, children were told that they would hear a label produced by another child, and that they should indicate whether this label adequately identified the target object. Three of the trials were far trials in which an unambiguously incorrect label was provided to the child (e.g. "table" to distinguish between a pug and a rabbit), three of the trials were close trials in which an unambiguously correct label was provided (e.g. "polar bear" to distinguish between a polar bear and a grizzly bear), and three were close trials in which an ambiguous label was provided (e.g. "dog" to distinguish between a husky and a dalmatian). Children responded verbally ("yes" it is adequate or "no" it is not) and an experimenter selected the corresponding button on the screen.

Procedure

After giving informed consent, adults read an introduction screen informing them that they would be doing an experiment in which they saw one or two pictures on the screen at a time. One of these objects would be in a blue box, and they should describe this object as best as they could by typing one or two words into a text box. At the start of each block, they would be told whether they would see one picture (isolation), or two pictures (close and far), and reminded that they should type in a description that would help another participant identify the target with a blue border. Adults then completed all blocks in random order.

After parents gave consent and children assented to participate, children were told that they would see pictures and be asked to describe them. Children then completed several warm-up trials to introduce them to the game. On the first trial, they saw a picture of a red apple. On the next trial, they saw a pear with a blue border around it and the same apple and were asked to describe the object in the blue box. Then, the experimenter indicated that they would close their eyes on the next trial so that they wouldn't see what children saw. On this trial, children saw a white mug with a blue border around it and a white plate. While the experimenter's eyes were closed, the child was encouraged to describe the object in the blue box so that the experimenter could identify it. Once children provided a description, the blue border disappeared, and the experimenter opened their eyes. The experimenter then identified the object based on children's descriptions, and revealed the blue box again to verify whether they selected the correct target.

The same procedure was repeated for three more trials with the following image pairs: a red purple ball and a plant; a white rose and a red rose; a round wooden table and a rectangular wooden table. The latter two trials contained object pairs where both objects shared a basic level category label (i.e., "flower" and "table"). These two trials served to subtly indicate to children that referential ambiguity is possible. If children did not uniquely identify the target (e.g., just saying "flower"), the experimenter would indicate that the child's expression was ambiguous, and the experimenter would not select the correct target. While these trials provided feedback about ambiguity to the children, the experimenter did not suggest alternate descriptions or examples of unambiguous expressions. After these warm-up trials, children were told that they would continue to play the labeling game and that their responses would be shown to another person who did not know which one was in the blue border. In contrast to adults, children always began with a block of familiar objects in either the close or far condition. They always completed the manipulation check block last.

Preprocessing

We cleaned the text input using a uniform rubric across the combined data set. First, we manually corrected typos and removed stop words (e.g., determiners like 'a', 'the'). Second, we lemmatized all entries to remove spurious differences between tenses and plurals of the same root form. The measures we use in analysis consider edit distance, which may be aritifically inflated due to word order (e.g., the edit distance between "running person" and "person running" would be high, despite conceptual and semantic similarity between the descriptions). As such, we manually cleaned the data to remove similar phrases or structures across descriptions (e.g., if a participant repeats "a person who is…" on every trial, we removed that phrase). We use the manually cleaned data for our analyses, but pre-cleaned, lemmatized data can be found on our OSF page (https://osf.io/vkug8/). Third, we removed spaces and collapsed multiple words together into a single token (e.g., 'German Shepherd' was tokenized to 'germanshepherd'). While adult participants typed in their own responses, children's responses were entered by an experimenter. When children's descriptions were overly long, experimenters prompted children to simplify them by asking, "Can you say that in one or two words?". Children were only prompted once, regardless of whether they simplified their expression.

2.3.2. Results

We focus on two primary hypotheses. First, could children be failing to take into account the referential context when deciding what to say, leading to more ambiguous or underinformative referring expressions? Second, could children simply have more uncertainty over possible acceptable labels for novel objects, making retrieval challenging (Cycowicz et al., 1997; Lachman et al., 1974)? These hypotheses are not mutually exclusive. Indeed, the corresponding mechanisms

– *pragmatic reasoning* and *lexical priors* – are both implicated in recent production models (Hawkins et al., 2022; e.g., Murthy et al., 2022).

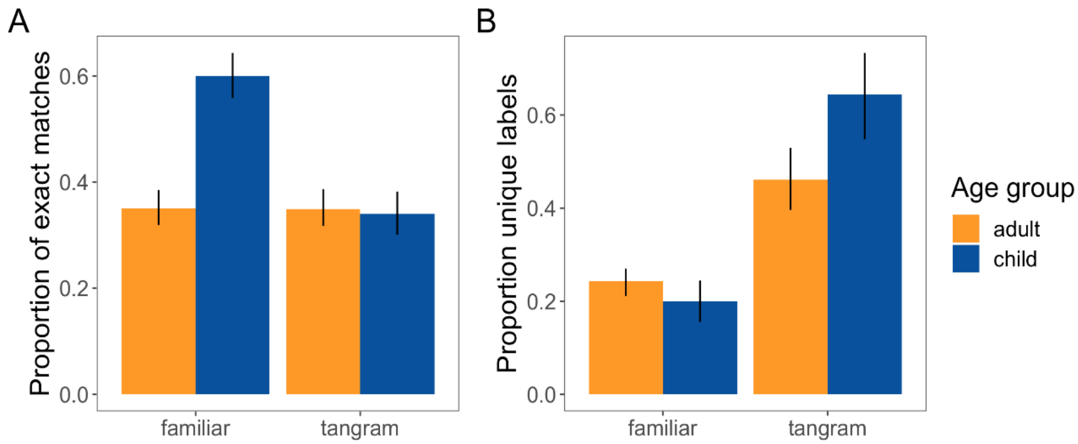

Figure 2.7 *(A) Context-sensitivity for children and adults. Children were twice as likely to give the same description for familiar objects than tangrams across contexts, while adults equally modulated their descriptions for both target types. (B) Nameability for children and adults. Children produced more variable labels for tangrams than adults. Error bars are 95% CIs.*

Children are differentially sensitive to referential context

To test our first hypothesis, we examine the extent to which the same participant produces different utterances across the *far* vs. *close* contexts. We begin by considering a simple 'exact match' criterion, coding an item as 1 if the participant used the same label for that item in both contexts and 0 if they used different labels. This criterion is conservative in the sense that it will miss a number of near- or partial-matches, giving a lower-bound for overlap. We model the binary variable of context overlap using a mixed-effects logistic regression. We include fixed effects for age cohort (child vs. adult) and target type (familiar vs. novel), as well as their interaction. To control for the fact that longer utterance strings are less likely to exactly match by chance, we also include a term for the average length of the close and far labels for that speaker and item. The most complex random effects structure that converged only included random intercepts at the participant level.

We found a significant interaction between age group and target type, $b = -1.09, z = -5.63, p < .001$ (see **Figure 2.7**). Although adults displayed similar rates of context-sensitivity for familiar and novel objects (familiar: $m = 0.35$, novel: $m = 0.35$), children were nearly twice as likely to provide the exact same label across contexts for a familiar object (familiar: $m = 0.34$; novel: $m = 0.60$). Similar results were obtained using 'softer' measures like edit distance, which is the number of edits required to turn one string into the other, $b = 0.18, z = 3.80, p < .001$. In other words, while adults appropriately modulated their utterances across contexts, children often produced the same description for familiar targets across contexts, even when it was not sufficient to distinguish the target from the distractor.

Familiar objects elicit less variable names

The previous section suggests that children have difficulty recognizing the referential ambiguity of a canonical label like "dog" in a context where there are multiple dogs. In other words, children's production may be constrained by limitations in their *pragmatic reasoning*. But why would insensitivity to pragmatic context be limited to familiar objects? One possibility is that children have a less stable prior over possible names for the novel tangram images due to lower 'codability' or 'nameability' (Hupet et al., 1991; Zettersten & Lupyan, 2020). Their variability across contexts could be driven less by pragmatic reasoning and more by "sampling variation" (Bonawitz et al., 2014; Denison et al., 2013).

We test this hypothesis by examining the distribution of labels at the population level. First, we hypothesize that both adults and children will produce a fairly narrow, high-agreement range of labels for familiar objects, yielding highly concentrated distributions. Second, we hypothesize that children will use a broader range of different labels for tangrams than adults, yielding a less concentrated distribution with less agreement among different children. We considered several measures of concentration, but we focus primarily on the proportion of unique labels to total labels, which is simple and interpretable.[6] For example, suppose that from a pool of forty participants, twenty said 'bird', ten said 'dancer', and the other ten said something unique. Then we would have $p_{uniq} = 12/40 = .3$. Meanwhile, if all forty participants chose different labels, we would have $p_{uniq} = 40/40 = 1$; and at the other extreme, if all forty participants chose the same label, we would have $p_{uniq} = 1/40 = 0.025$.

Because population agreement metrics necessarily aggregate over individual participants for each target, we construct our mixed-effects regression model at the item level. Given our findings of differential context-sensitivity in the previous section, we limit this analysis to the 'far' condition where the distribution of adult and children labels are more comparable (aggregating across close and far yields qualitatively similar results). We predict agreement as a function of age group (adult vs. child) and target type (familiar vs. tangram), including random intercepts at the target level. First, we observe a main effect of target type, with less agreement on tangram labels for all participants, $b = -0.16$, $t(15.08) = -7.98$, $p < .001$. This is in line with our hypothesis that agreement would be higher for familiar objects with commonly-known canonical labels. Importantly, however, we also find a significant interaction with age group, $b = 0.06$,

[6] In principle, the most appropriate metric of spread across labels would be the information theoretic quantity of entropy: $H(X) = \sum_i p_i \, log p_i$. However, our empirical distributions are highly sparse, with many labels appearing only once. Our estimates of entropy are therefore somewhat sensitive to the choice of statistical estimator (e.g. how much to regularize with pseudo-counts) while being less numerically interpretable. However, we find a similar qualitative effects of interest for reasonable choices, \$b=\$0.03, t(13.52)= 3.47, p = 0.00 using the Schurmann-Grassberger estimator. Alternative metrics include "modal agreement" (Brodeur et al., 2010, 2014), the proportion of participants that produce the most common label, and Simpson's diversity index (Majid et al., 2018; Majid & Burenhult, 2014; Simpson, 1949), which can be interpreted as the probability that two independently sampled labels will match.

$t(13.88) = 5.79, p < .001$. The labels produced by different children in our sample agree with one another about as much as adults' agree for familiar targets ($m = 0.20$ for children and $m = 0.24$ for adults). However, children as a group produce a much more variable set of labels for novel tangrams than adults do ($m = 0.64$ for children and $m = 0.46$ for adults).

Children's understanding of referential ambiguity predicts production

Children are generally less sensitive to referential context than adults, but what contributes to these differences? Given that children are more likely to produce the same label across contexts for familiar objects, it is unlikely that vocabulary knowledge accounts for most of the difference between children and adults. One factor that may influence children's performance is the recognition of referential ambiguity. That is, children who understand that ambiguous referring expressions are not sufficiently informative for a partner should show more context sensitivity in our production task. To assess that children understand the goal of the task, and to aid in our interpretation of any potentially ambiguous responses that children give in earlier trials, we included a manipulation check condition for children. We first report children's responses to all three check trial types. Then, we turn to analyze how responses to ambiguous check trials (e.g., "dog" to distinguish between a husky and a dalmatian) predicts children's context sensitivity on the production task.

For all check trials, children were asked to respond yes or no as to whether a provided label uniquely identifies one picture out of two. If children are sensitive to informativity of referential expressions, they should respond yes to close trials with *unambiguous* and correct labels (e.g., "red flower" which distinguishes a rose from a lily), no to far trials where the provided labels were unambiguous and *incorrect* (e.g., "table" which fails to distinguish between a pug and a rabbit), and also no to close trials with *ambiguous* expressions (e.g., "dog", which fails to distinguish between a husky and a dalmatian). First, we found an overall effect of age, where children's accuracy on the check trials increased in older cohorts, $b = 0.12$, $t(166.65) = 4.29$, $p < .001$. Second, we found that children typically responded as expected to the unambiguous trials. That is, they typically responded no to the incorrect trial (89.64% correct), and yes to the unambiguous trial (97.07% correct). Ambiguous trials were more challenging. Across all ages, children made more errors on ambiguous trials (56.06% correct) relative to incorrect trials, $b = 0.67$, $t(116.00) = 2.87$, $p = .005$) and unambiguous trials, $b = 1.07$, $t(116.00) = 4.57$, $p < .001$. These findings suggest that children, regardless of age, understood that that goal of the task was to describe a target picture (and not, for instance, to simply say anything).

To assess whether recognition of referential ambiguity correlated with performance on the production task, we focused specifically on children's performance on the ambiguous check trials. On these trials, children were asked whether a given description was sufficient for distinguishing a target from its distractor. We predicted that children who could recognize ambiguity on these trials would also be more likely to produce informative descriptions. We fit a mixed-effects model

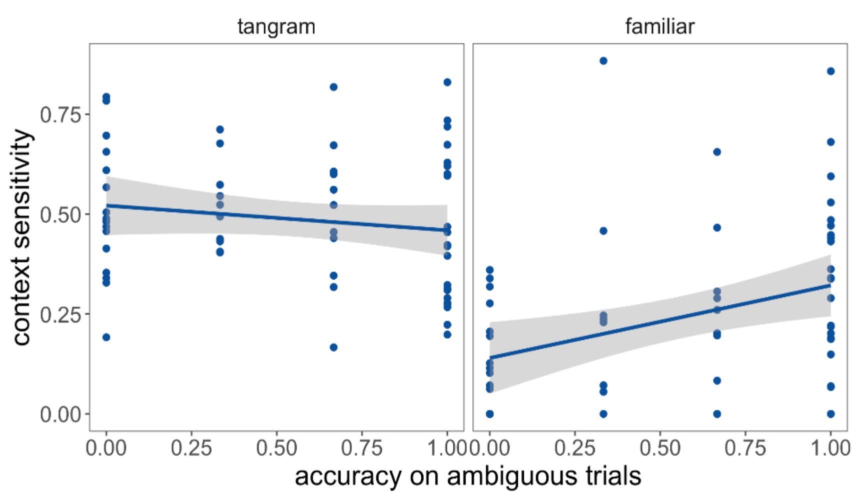

Figure 2.8 *Children's performance on ambiguous manipulation check trials predicted their context sensitivity (measured by exact matches in descriptions between close and far contexts) for familiar objects, but not for tangrams. Note that there were only three ambiguous manipulation check trials, thus only four possible accuracy rates*

predicting children's context sensitivity, as measured by 'exact match' overlap, from their performance on the ambiguous trials and target type (familiar vs. tangram), with random effects of participant and age. We found a significant interaction between accuracy on the ambiguous trials and target type, such that children who were more accurate on the ambiguous check trials were more likely to modulate their descriptions based on context for familiar items, $b = 0.24, t(112.87) = 2.98, p = .003$. In other words, children who were more sensitive to referential ambiguity were less likely to produce the same descriptions for familiar items across close and far contexts (**Figure 2.8**).

2.3.3. Discussion

In Experiment 2.3, we asked whether difficulty in *production* may explain why young children struggle to form referential pacts with peers. Comparing adult and child responses on our production task revealed that children were not as sensitive to referential context, providing descriptions that did not always distinguish the intended referent from its distractor (e.g., saying "table" when there were both a dining table and a picnic table). Importantly, this effect was found on trials containing familiar items that children are likely to have accessible labels for, so these results were unlikely to be explained by simple vocabulary constraints.

Why might children struggle to modulate their descriptions for familiar objects? One possibility is that the basic level category label of a familiar object (e.g., "table") is too salient, and children have difficulty suppressing this label in favor of a more informative description. While we cannot infer whether children were considering alternative descriptions during our task, we did find that children, compared to adults, produced less variable descriptions for familiar objects, concentrating on the same canonical labels. Further studies probing the alternatives that children consider, or testing the salience of various object labels, could provide a stronger test of this possibility.

Another possibility is that children are not sensitive to referential ambiguity. Indeed, prior work shows that children may not know which expressions are more informative (e.g., Beal & Flavell, 1982; Robinson & Robinson, 1977). Our experiment provides further evidence for a direct link between *recognizing* referential ambiguity and *generating* appropriately informative expressions based on context. Children who were able to recognize that a given ambiguous expressions would be not helpful were more likely to show context sensitivity in their own ability to generate descriptions for familiar objects. By the same token, failure to recognize that a description is ambiguous would likely lead to communicative breakdown—young children may not know to ask for clarification when necessary.

If children do not reliably produce (or identify) helpful descriptions, how did they succeed in Experiment 2.1? The success we observed in parent-child pairs may be attributed to the supportive context that parents provided. Interactive behaviors such as requesting more specific information or following up with specific questions could provide children with feedback that their expressions are not working, and suggest stronger possibilities. Communicating with parents, who have specific knowledge about their children, as well as abundant experience interacting with their children, may be especially helpful for facilitating young children's referential abilities. Further research comparing parents to other adults (who have less specific knowledge about a particular child) will provide insight into whether parents play a unique role above and beyond adults in general.

2.4. General Discussion

Successful communication crucially depends on the ability to establish common ground and coordinate on conversational pacts with partners. A deep literature of developmental work going back to Krauss and Glucksberg (1977) has revealed startling failures to coordinate even among older children who are otherwise highly competent users of language. These failures have presented a persistent puzzle for theories of social and cognitive development, raising a number of possible explanatory mechanisms. Here, we turned to parent-child interactions for insight. Surprisingly, even young children were able to successfully form conversational pacts with parents. Whatever obstacles arise in peer communication are to some extent resolved through interactive parental scaffolding; parents of younger children engaged in more conversational turns, asked more questions, produced longer referential expressions, and adopted less complex labels. We then conducted two additional studies to disentangle the contribution of comprehension and production processes, determining that *production* is likely to present the root problem that parental scaffolding is constructed to overcome.

Our studies contribute to the debate over why children may struggle to form referential pacts with peers. One classic family of explanations implicates theory of mind development: children may fail to take their partner's perspective into account and stubbornly produce idiosyncratic descriptions that are not meaningful to anyone else (Krauss & Glucksberg, 1977). Another family of explanations implicates sensitivity to informativeness in context: children may be focused on generating an appropriate standalone label for the target while neglecting to consider whether it might also apply to distractors and lead to confusion (Speer, 1984). Our study introduces a third possibility. Young children may be unable to generate a good enough label on their own, due to weak lexical priors for what to call the most informative features of unfamiliar objects.

Once presented with a good-enough label from a parent, however, they are flexible enough to readily adopt and use it themselves.

An important limitation of our study is our reliance on existing empirical data documenting failures in child-child communication, rather than more directly comparable data replicating these failures in the same paradigm we used to study parent-child communication. It is possible that our paradigm differs in some important parameter from those used in prior work and pairs of children would, in fact, succeed at forming conversational pacts in our variant. For example, our task uses contexts containing only two images per trial, whereas classic studies used contexts of six or more. This choice could reduce cognitive load in our variant, allowing child-child pairs to perform as well as our parent-child pairs. At the same time, there are several reasons to believe that pairs of children would still struggle in our variant, including the production failures we independently observed in Experiment 2.3 and the observation than younger children relied more heavily on parent-generated descriptions in Experiment 2.1. Further experiments are necessary to replicate classic child-child phenomena under our paradigm, or replicate our parent-child phenomena under classic paradigms.

Taken together, these studies contribute to the growing understanding of parental scaffolding in language development more broadly. As children encounter novel concepts and situations they do not yet have words to describe, they must be particularly attuned to the distribution of labels that appear to be acceptable to competent speakers. Beyond the benefits of child-directed speech (ManyBabies Consortium, 2020) and careful tailoring of labels to the child's vocabulary (Leung et al., 2021), adults may spontaneously adopt some of the higher-order discourse strategies examined in training studies (Matthews et al., 2007). These more interactive forms of scaffolding go beyond a single utterance to draw attention to sources of ambiguity and expose the need to actively check for mutual understanding. The everyday process of parents and children trying to understand one another may in this way not only scaffold the formation of conversational pacts, but of communicative development more broadly.

Chapter 3. A Closer Look at Communicative Perspective-Taking: Children's Estimates of Other's Specific Knowledge

In the Introduction and Chapters 1-2, I have conceptualized communication as an *adaptive* process. Successful communication relies on making adjustments based on our listeners, and Chapters 1 and 2 provide evidence that parents sensitively adapt their speech according to their children's developmental level. In Chapter 2, I began to explore possible challenges that young children face when communicating with others, and suggest that children may have particular difficulty with producing informative referring expressions. In Chapter 3, I examine another aspect related to children's communicative-perspective-taking. Namely, the ability to make accurate estimates about another person's knowledge.

Representing another person's vocabulary knowledge is potentially important even for daily conversation: Imagine visiting the zoo with your friend and their 2-year-old. As you walk by the peacocks, you hear your friend say, "Do you see those blue birds?" Immediately, you know that your friend is talking to their child and not you. If they were talking to you, saying "peacock" would be perfectly clear; however, "blue bird" might be a better description for a child who has never seen a peacock before. Even when talking about the same object, we use different words depending on what we think our conversational partners know and don't know. In service of effective communication, we can leverage our models of other people's knowledge to design speech that is suited for our partner.

The fluency of our everyday conversations depends on exactly this kind of adaptation. A large body of work has documented the variety of ways in which adults take their communicative partners' knowledge into account (e.g., Clark & Murphy, 1982; Brown-Schmidt & Hanna, 2011). For example, when re-telling a story to someone who has heard it before, adults reduce the amount of information they give, but not when re-telling the story to a new partner (Galati & Brennan, 2010). Adults can adapt even to partners who are quite different from them, as in the case of parents and their children. Parents leverage their fine-grained models of their children's vocabularies and use this knowledge in spontaneous communication, as demonstrated in Chapter 2 (e.g., using "blue bird" to describe a peacock; Leung, et al., 2021). Taken together, these studies show that adults tailor their speech according to their conversational partner's knowledge, their previous interactions, and general metalinguistic knowledge.

In comparison, children may be less flexible in their ability to adapt to communicative partners. In studies that explicitly highlight their conversational partner's knowledge, children can use this knowledge to guide their word choices (Baer & Friedman, 2018). However, other studies have shown that children struggle to adapt to their communicative partners in more naturalistic conversational settings (e.g., Krauss & Glucksberg, 1977). In children's everyday interactions, effective listener design requires them to spontaneously reason about their partner's knowledge and adjust their speech accordingly. In this chapter, I directly probe one crucial component of listener design: the ability to infer a partner's vocabulary knowledge.

Adults can make graded and surprisingly accurate relative estimates of when a word is generally learned. Kuperman et al. (2012) asked adult participants to report the age at which they understood a given word and obtained judgments for 30,000 English words. These judgments were then directly compared with age of acquisition data, i.e., the typical age that a given word is actually learned (hereafter referred to as AoA). While adults typically overestimate the *absolute age* at which they learned a given word, the estimated *order* in which words are acquired is intact (Kuperman et al., 2012). This metalinguistic knowledge could allow adults to make reasonable inferences about what different individuals (e.g., a young child) might know and adapt their speech accordingly.

Can children use this same kind of information to predict what words a younger child might know? Very few studies have investigated children's estimates of other people's specific vocabulary knowledge, but there is some evidence suggesting that even young children may have surprising metalinguistic knowledge (Walley & Metsala, 1992). Children as young as 5 years old accurately estimated the order in which they acquired words. More incredibly, they were sensitive to the order in which they might acquire new words (Walley & Metsala, 1992). That is, when asked when they expect to learn an unknown word, children might be able to appeal to their own lexical and metalinguistic knowledge to make sensible judgments. Even so, reasoning about *another person's* specific lexical knowledge could be difficult for young children. Children often over-attribute knowledge to others, especially knowledge they themselves already have (Birch & Bloom, 2003; Ghrear et al., 2021). This bias to over-attribute knowledge could hinder children's ability to reason about a younger child's knowledge: they may have trouble inhibiting their own vocabulary knowledge to correctly identify something that a younger child does *not* know.

However, there is evidence that even preschool age children can make nonegocentric knowledge judgments in some situations. When asked about variety of general knowledge skills, young children are sensitive to the fact that children and adults possess different levels of knowledge (Fitneva, 2010; Taylor et al., 1991). In a series of experiments, Taylor et al. (1991) found that 4- and 5-year-old children recognized what facts are known by adults but not children (e.g., What a certain plant looks like), as well as facts that children may know but babies may not (e.g., What a certain shape is called). Children do not simply believe that adults always know more than children—they are sensitive to the types of knowledge that each group may possess. In

Fitneva's (2010) study, 4- to 6-year-old children attribute certain child-specific knowledge to children only (e.g., names of cartoon characters), indicating that they do not have a general belief that adults are necessarily more knowledgeable than children in all circumstances. While these studies sometimes include vocabulary items (e.g., Taylor et al., 1991), they test whether children make broad distinctions between different people's knowledge, such as an infant not knowing any words, a child knowing simple words (e.g., *happy*), and an adult knowing complex words (e.g., *hypochondriac*).

Why is it important to examine children's ability to estimate other people's specific lexical knowledge? The broader question I hope to address in this chapter (and this dissertation) is how children become fluent conversationalists like us. Because communication relies on adaptation, it is important to understand how children navigate the process of adapting to others. Further, many studies indicate that children struggle to coordinate with others (usually peers) in referential communication (e.g., Krauss & Glucksberg, 1977). To what extent to children's difficulties come from errors in estimating others' knowledge? In Chapter 2, I found that children may have general cognitive and linguistic limitations that lead to *production* difficulties: young children have a hard time generating helpful descriptions of objects. Beyond their own limitations, however, it is possible that children have additional difficulty making fine-grained judgments about vocabulary knowledge. They may realize that adults are more likely to know words like *hypochondriac* (Taylor et al., 1991), but they may not distinguish between words that are generally acquired relatively close together: for example, children may not recognize that a 4-year-old is more likely than a 2-year-old to know the word *lobster*. If children struggle to identify the lexical items that someone else knows or does not know, communicative adaptation may not occur.

If children do make sensible estimates of others' vocabulary knowledge, can they use these estimates to adapt to others? One possibility is that, like adults, children make use of their models of others to design speech for effective communication. Indeed, one study by Baer and Friedman (2018) found that children as young as 5 modify the types of information they provide based on their listener's knowledge state. When talking about a familiar object, such as an umbrella, children provided general information (e.g., "umbrellas are used in the rain") to naïve listeners, but specific information (e.g., "this umbrella has polka dots") to knowledgeable listeners. On the other hand, children appear to have difficulty coordinating multiple perspectives in communication (e.g., Krauss & Glucksberg), and may not spontaneously adapt to others without explicit feedback.

Thus far I have discussed two skills relevant to communicative adaptation: 1) the ability to estimate another person's specific knowledge; 2) the ability to leverage models of others' knowledge to design suitable speech. Do young children have each of these skills, and how do these skills development across early and middle childhood? In this chapter, I first ask whether children can infer another child's specific vocabulary knowledge to make word-level predictions consistent with normative AoA. One study suggests that children as young as 5 can accurately estimate the age and order in which they learned a variety of words (Walley & Metsala, 1992), but can they reason about other children's vocabulary knowledge? In Experiment 1, 4- to 8-year-old children were introduced to a younger fictional child, and asked to make judgments about the target child's knowledge of various familiar words. Even 4-year-old children made judgments that matched the estimated order of acquisition, such that they judged the fictional child to be more likely to know early-acquired words and less likely to know late-acquired words. Compared to younger children, older children's judgments more reliably recovered the order of acquisition. In Experiment 2, which is ongoing, 4- to 8-year-old children are asked to communicate with a younger fictional child. The key question is whether children will spontaneously adapt their speech based on beliefs about the younger child's lexical knowledge. Our current exploratory analysis of a partial sample indicates that children may not spontaneously adapt to the younger child, even when the younger child's lexical knowledge is explicitly provided.

3.1. Experiment 3.1: Children's estimates of other's vocabulary knowledge

3.1.1. Methods

Stimuli

To create a coherent game that would be enjoyable for children, we selected stimuli from a single domain (animals). Our stimuli consisted of 15 animal words, along with corresponding images of each animal. We pulled all animal images (n = 45) from a normed image set (Rossion & Pourtois, 2004; recoloring of Snodgrass & Vanderwart, 1980). To ensure our stimuli spanned a range of ages of acquisition (AoAs), we ranked the animal words from earliest to latest AoA, using adult estimates from Kuperman et al. (2012), and split the words into five bins. In order to select animal images that are recognizable and typically identified by a single name, we chose the three animals from each AoA bin with the highest naming agreement according to a naming task with children (Cycowicz, Friedman, Rothstein, & Snodgrass, 1997).

The resulting animal words, ordered by estimated AoA, were *dog, duck, cat, pig, fish, turtle, zebra, elephant, snake, penguin, gorilla, owl, raccoon, leopard,* and *lobster*. Although adult AoA estimates for these words range from 2.5 to 7.5 years old (Kuperman et al., 2012), all of these animal words are generally acquired by age 3 according to parent-reported estimates of children's vocabulary knowledge (Frank et al., 2017). Because the youngest children in our study were 4 years old, we expected all participants to know these animal words. Children were also provided with the canonical labels for all animals at the beginning of the study, in order to ensure that all children knew what each animal was called.

Participants

We pre-registered a planned sample of 60 children ages 4-8, with 12 children per year-wise age group. Due to over recruitment, our final sample included 62 children (12 4-year-olds, 13 5-year-olds, 13 6-year-olds, 12 7-year-olds, 12 8-yearolds). Based on a pre-registered exclusion criterion, children who failed to answer all of the questions were excluded and replaced (an additional 6 children). Families were recruited online, primarily through a US University database of families who have expressed interest in doing research or previously participated. Children completed this study over Zoom, interacting with a live experimenter who navigated a slide-style, animated Qualtrics survey.

A separate sample of 30 adults were recruited via Amazon Mechanical Turk. The adult sample provides a simple test that our task elicits robust inferences about the target child's lexical knowledge, and that these inferences correspond to extant AoA data. Adult participants completed the same task using Qualtrics, with minor modifications as described below.

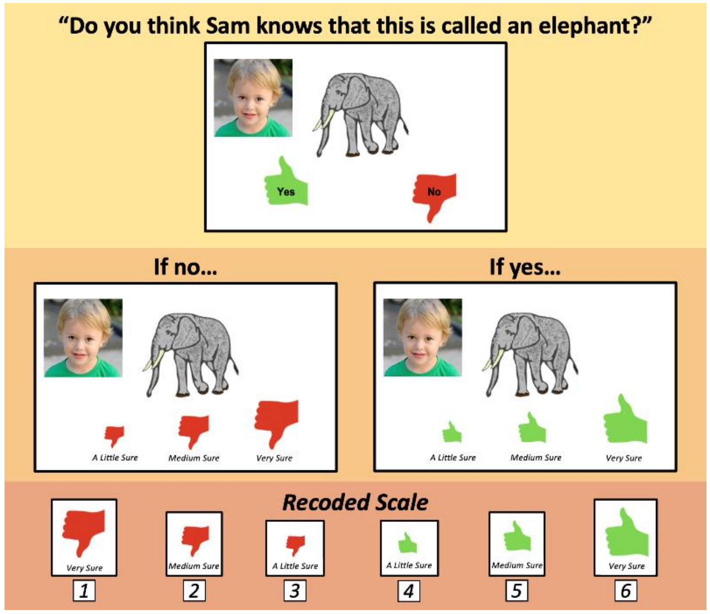

Figure 3.1 *Trial structure. The experimenter labeled the animal, then asked the child "Do you think Sam knows that this is called an elephant?" Based on their response, children were then asked to provide a confidence judgment on a 3-point scale (a little sure, medium sure, very sure). Confidence judgments were recoded into the 6-point scale shown here.*

Procedure

Introduction. Children were shown a picture of a child named "Sam" (see ***Figure 3.1***). Children were anchored to Sam's knowledge of various familiar skills, specifically some skills that Sam has acquired (e.g., coloring), and some that Sam has not yet acquired (e.g., reading). Children were then specifically anchored to Sam's possible word knowledge in a nonanimal domain. They were given an example of one word Sam knows (*car*), and one word that Sam does not know (*piano*). This introduction was intended to ensure that children understand there are things Sam does *not* know yet (even things children themselves likely know, such as how to read).

Trial structure. On each trial, children were shown a drawing of a familiar object or animal (Rossion & Pourtois, 2004). The experimenter first labeled the object (e.g., "Look, it's [an elephant]!"), and then asked about the target child's knowledge (e.g., "Do you think Sam knows that this is called [an elephant]? Yes or no?"). Based on their response, children were then asked a contingent follow-up question: "How sure are you that Sam [knows/doesn't know] that this is called [an elephant]–a little sure, medium sure, or very sure?" All questions were presented with accompanying pictures of thumbs [up/down] of varying size (see ***Figure 3.1***). Children as young as 3 are able to engage in uncertainty monitoring and report their confidence, although these skills do develop in the preschool years (Lyons & Ghetti, 2011). Children's responses to these two items were recoded onto a 1-6 scale from 1–*very sure Sam doesn't know* to 6–*very sure Sam knows* (***Figure 3.1***). Our two-step question structure allowed us to collect a gradient response while maintaining simplicity, as young children may struggle to comprehend a 6-point scale.

The experimenter provided no evaluative feedback on any trials, but did offer consistent neutral feedback (e.g., repeating the child's answer or saying "Okay!"). When a child failed to respond within about 5 seconds or offered a noncanonical response (e.g., saying "Maybe"), the experimenter acknowledged the child's answer and then repeated the question with the possible responses. If a child did not answer after the question was repeated, the experimenter moved on and marked the trial as no response. These were considered "incomplete" sessions and these participants were not included in the final sample.

Familiarization trials. Children first completed two nonanimal familiarization trials, one for an early-acquired word (ball) and one for a late-acquired word (artichoke). These trials followed the trial structure described above and were intended to help familiarize children with the structure of the questions and scales. These trials were always asked first and in a fixed order.

Animal trials. Children were then shown 15 trials of the same form (see example trial in ***Figure 3.1***). For the 15 animal trials, trial order was randomized across participants to control for any potential order effects in children's responses.

Explanations. After completing the final animal trial, children were asked an open-ended explanation question about their final judgment (e.g., “Why do you think Sam [knows/doesn’t know] that this is called [an elephant]?”). Because the trial order was randomized, the explanations concerned different animal words across participants.

Final check questions. Children were asked two questions about Sam’s skill knowledge, one early-acquired skill (going up and down stairs) and one very late-acquired skill (driving a car). These questions again followed the general trial structure described above. The skill knowledge items were included as an additional check that children at all ages were able to use the scale appropriately, in case young children failed to differentiate animal words based on AoA. Lastly, children were asked to report how old they thought Sam was. This question was intended to assess another aspect of children’s belief about Sam. Sam’s photo and skill knowledge were intended to indicate toddlerhood.

Adult procedure. Adult participants completed a minimally adapted version of the same task online via Qualtrics. Unlike children, adults were simply presented with the full 6-point scale (1–*very sure Sam doesn’t know* to 6–*very sure Sam does know*). Additionally, the task was administered asynchronously, so adult participants did not interact with an experimenter or receive any feedback during the task. Otherwise, the adult task was identical to the child task described above.

3.1.2. Results

Familiarization trials

Two familiarization items (*ball* and *artichoke*) were included to help children get accustomed to the general trial structure. We report children’s responses on these familiarization items here. We used a mixed effects model using the `lme4` package in R (Bates, Mächler, Bolker,

& Walker, 2015), predicting children's knowledge judgments from the item with a random effect of participant.

Overall, children were significantly more likely to report that Sam knows the word *ball* (*mean* = 4.65) than that Sam knows the word *artichoke* (*mean* = 1.87, $b = 2.77$, $t = 10.6$, $p < .01$). Analyzing judgments separately for each age group, 4-year-olds did not significantly differentiate the two familiarization items ($b = 0.17$, $t = 0.21$, $p = .83$). All other age groups significantly differentiated the two familiarization items (*ps* < 0.05).

Skill knowledge

As an initial check that children at all ages were able to use the scale appropriately and infer knowledge in an easier case, we included two questions about the target child's skill knowledge. Note that the two skill items (going up and down stairs and driving a car) are in line with children's own knowledge. That is, children should be able to answer these questions appropriately even if they are reasoning egocentrically about their own knowledge.

Overall, children differentiated the target child's skill knowledge on these two items. We used a similar mixed effects structure predicting children's knowledge judgments from the item with a random effect of participant. Children were significantly more likely to report that the target child knows how to go up and down stairs (*mean* = 4.1) than that the child knows how to drive a car (*mean* = 1.4, $b = 2.69$, $t = 10.29$, $p < .01$). Analyzing judgments separately for each age group, even 4-year-olds significantly differentiated the two skill items ($b = 2.33$, $t = 4.31$, $p < .01$).

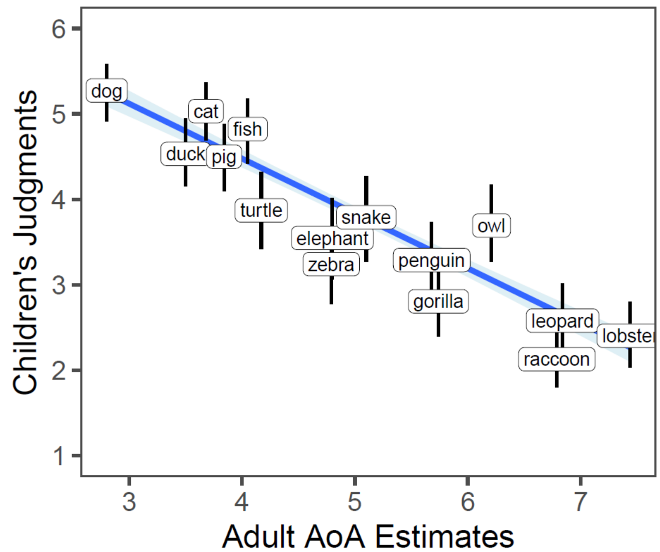

Figure 3.2 *Comparing adult AoA estimates (in years, taken from Kuperman et al., 2012) and children's judgments on our 6-point scale (1 = very sure Sam doesn't know; 6 = very sure Sam knows). The black lines show 95% confidence intervals for each item. The shaded region shows the confidence interval based on a linear regression estimated from the raw data.*

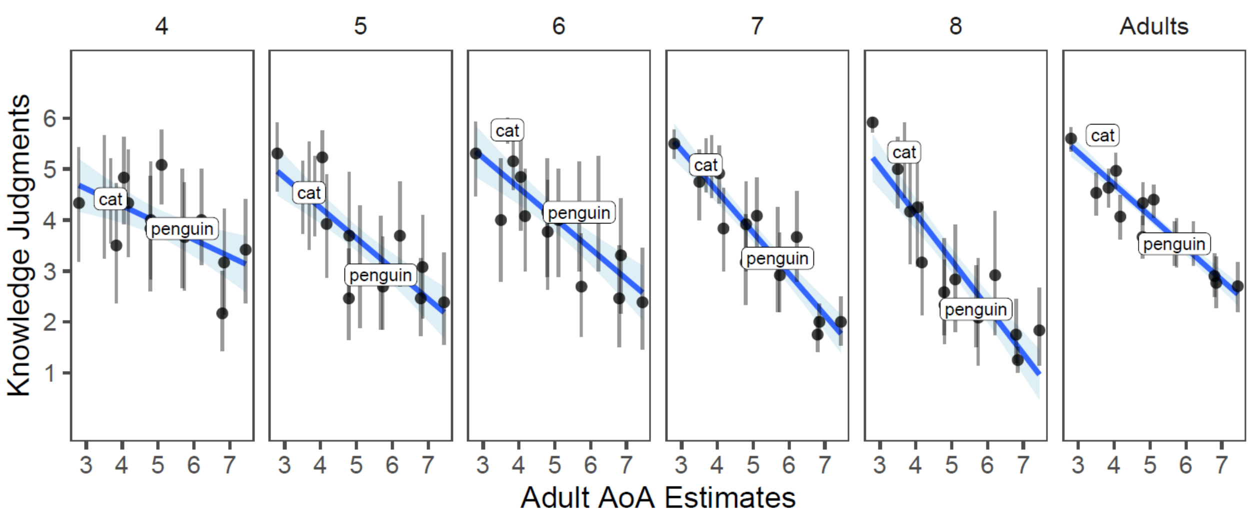

Figure 3.3 *Children's and adults' judgements about the target child's word knowledge across development, compared with adult AoA estimates (in years, taken from Kuperman et al., 2012). Each point represents 1 of the 15 word items, with black lines showing 95% percent confidence intervals for each item. The shaded region shows the confidence interval based on a linear regression estimated from the raw data.*

Judgments of vocabulary knowledge

Our primary analyses compare knowledge judgments on our 6-point scale to AoA estimates from adults (taken from Kuperman et al., 2012). Data were analyzed using a preregistered mixed effects model. We predicted knowledge judgments from adult AoA estimates, including random effects for participant and word.

We expected that overall, children's judgments would recover the ordinal shape of age of acquisition data for these items. That is, children would infer that the target child is most likely to know early-acquired words, and least likely to know late-acquired words. As a result, we expected a negative relationship between judgments of the target child's lexical knowledge and adult AoA estimates.

First, analyzing adults responses on our task, we saw the predicted negative correlation between AoA and adults' judgments of the target child's knowledge (***Figure 3.3***, $b = -0.63$, $t = -8.71$, $p < .01$). This confirmed that our task elicited reliable predictions from adults, and that adults' inferences about the target child's knowledge match predictions from AoA estimation tasks (Kuperman et al., 2012).

Do children's judgments about another child's vocabulary knowledge also reflect a sensitivity to which words are learned earlier or later? Overall, we found a significant negative correlation between AoA and children's judgments ($b = -0.65$, $t = -8.29$, $p < .001$). As a group, children were more confident that the target child would know an early-acquired word (e.g., *dog*), and also more confident that the target child would not know a late-acquired word (e.g., *lobster*, see ***Figure 3.2***).

We then asked whether children develop sensitivity to Sam's vocabulary knowledge, with older children's judgments recovering word-level AoA data more closely. We used the same mixed effects model but included an effect of age and an interaction between AoA and age. We again found a reliable main effect of AoA (b = -0.65, t = -8.29, p < .001), a main effect of age (b = 0.55, t = 3.68, p < .001) and a significant interaction between the two (b = -0.14, t = -5.1, p < .001). As predicted, older children's judgments were more adult-like, such that they more robustly reflected adult estimates of the order of acquisition (***Figure 3.3***).

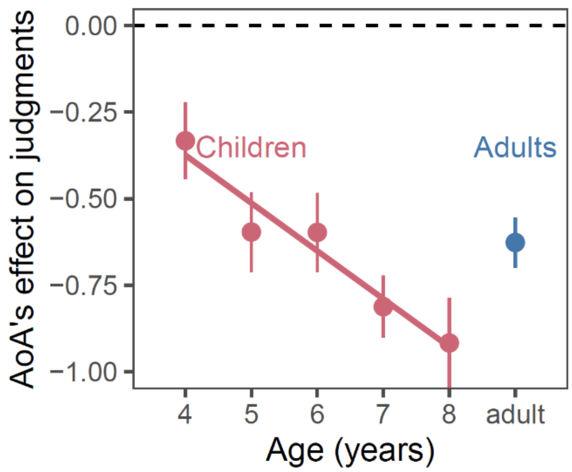

Figure 3.4 *Coefficient estimates of the effect of age of acquisition on children's and adults' knowledge judgments. Points indicate means, error bars indicate 1 standard deviation.*

To test the robustness of children's intuition at each age, we ran the model separately for each pre-determined year-wise age group (***Figure 3.4***). We found a significant negative correlation between AoA and children's judgments at all age groups (with the smallest effect in 4-year-olds: b = -0.33, t = -3.01, p = .01). That is, even 4-year-old children judged that late-acquired animal words were less likely to be known by the target child. Interestingly, judgments from the older two age groups of children were more closely correlated to data from Kuperman et al. (2012) than were adult participants' judgments (***Figure 3.4***). This appeared to be primarily driven by a greater willingness to judge Sam as moderately or very unlikely to know late-learned animal words, whereas adults were less sure about these same judgments (***Figure 3.3***). We return to this finding in the Discussion.

Target child age

At the end of the study, participants were asked to guess the target child's age. While the familiarization phase included information about the child's language and skill knowledge, no age was explicitly given. Looking at children's responses, the median response was that the target child was 3 years old. Looking at adult's responses, the median response was that the target child was 4 years old.

Explanations

As an exploratory analysis, we examined the reasons children gave for why the target child would or would not know a given word. While children sometimes offered spontaneous explanations throughout the study, our analysis focused on the explanations elicited after the final animal trial. The explanations were divided into 6 categories: *Language, Experience, Location, Age, Unsure,* and *Other*.

Language includes explanations that explicitly appealed tolanguage properties. *Experience* includes explanations that appealed to real-world experience with the referent. *Location* includes explanations that specifically referenced a particular place the animal is associated with. *Age* includes explanations that referenced a particular age or general age group. Any child that failed to answer the explanation question or expressed ignorance was coded as giving an explanation of *Unsure*. Explanations that did not fall into any of the above categories were coded as *Other*. Note that coding was not mutually exclusive, so explanations could be coded as including multiple categories. See Table 1 for examples of each category. **Figure 3.5** shows the proportion of children who gave each type of explanation.

To understand how children's explanations may change over development, we divided participants into older (6-8 years old) and younger children (4-5 years old). Unsurprisingly, *Unsure* explanations were much more common in younger children (44%) when compared to older children (8.33%). *Language* explanations were used by the highest proportion of children overall (27.27%). Do older children account for all of those explanations? Although more of the older children appealed to *Language* explanations (30.56%), these explanations were also common in younger children (24%). Thus, while young children were more likely to offer no explanation, the explanations they did offer seemed to rely on factors similar to older children's explanations.

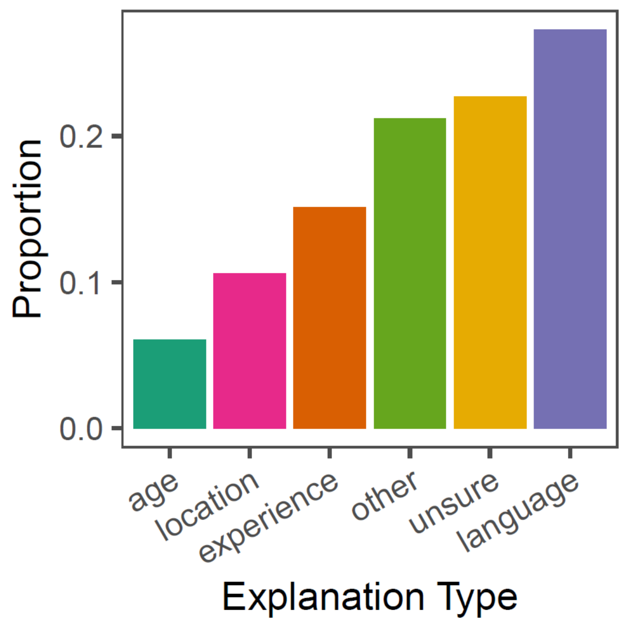

Figure 3.5 *Children's explanations for why they think the target child knew or didn't know an animal word. Categories are not mutually exclusive.*

3.1.3. Discussion

Young children are capable of inferring others' general knowledge, but are they also sensitive to another person's specific knowledge? We asked 4- to 8-year-old children to estimate another child's knowledge of lexical items and found that children as young as 4 are sensitive to a younger child's vocabulary knowledge.

Our findings highlight that young children have robust metalinguistic knowledge (Walley & Metsala, 1992), and can use that knowledge to make highly specific inferences about other people's vocabularies. The animal words used in our study are generally learned within a 6-month period, yet young children still distinguished early-acquired words from late-acquired words in this set. Prior studies have shown that children are sensitive to broad differences in vocabulary knowledge of infants, young children, and adults (Taylor et al., 1991). Our study further demonstrates that children readily make specific, word-level predictions about the language knowledge of another child.

Surprisingly, children's judgments of another child's knowledge did not just approach adults' judgments over development. Compared to adults, the oldest children in our study gave more accurate judgments as measured by their correlation with an external measure of age of acquisition (Kuperman, et al., 2012). While older children and adults were both highly confident that a young child would know early-acquired animals, adults were less confident that a young child would not know the late-acquired animals. It is possible that children are more accurate in their judgments because they better remember learning the animal words (see also Walley & Metsala, 1992), but our findings do not support this account, since older children were more accurate than younger children. Alternatively, adults' caution in asserting that the target child did not know the late-acquired animals could have reflected their difficulty in estimating the child's age. In line with this account, adults judged Sam to be a year older than children did on average. In future work, we plan to explore these possibilities by asking children and adults to make vocabulary judgments about children of multiple ages.

How are children in our study making estimates about other people's knowledge? Children's own explanations suggest that they use various cues to make their estimates. Overall, language-related explanations were most common, and even preschool age children appealed to this explanation. However, such explanations are difficult to interpret, and the mechanisms underlying children's knowledge estimates are outside the scope of the current study. Future work should more directly probe the features underlying this inference—to see if children are relying on their own uncertainty, word length (and other linguistic information), features of the referent itself, or other cues.

3.2. Experiment 3.2: Do children spontaneously adapt to a conversational partner?

In Experiment 3.1, we found that children as young as 4 are sensitive to other people's vocabulary knowledge, judging that a younger child would be more likely to know early-acquired animal words, and less likely to know late-acquired ones. While there is significant development in the accuracy of estimates between ages 4 and 8, our findings nonetheless indicate that young children have surprisingly fine-grained and reliable representations of other people's lexical knowledge. Can they leverage these representations during communicative instances to adapt to their conversational partners?

At least one study has shown that by age 5, children selectively talk about general or specific characteristics of an object based on their partner's knowledge state (Baer & Friedman, 2018). In Baer and Friedman's study, experimenters told children about their partner's knowledge before each trial, thus removing the need for real-time inferences to be made. Further, children in that study were not asked to distinguish between differences in specific lexical knowledge, but rather knowledge about categories or object groups. In Experiment 3.2, I ask whether children can use their real-time inferences of another person's vocabulary knowledge to adapt their descriptions of familiar animals.

3.2.1. Methods

Stimuli

The stimuli for Experiment 3.2 are identical to those used in Experiment 3.1, except that *fish* and *leopard* were removed. The reason for removing two items was to shorten the number of trials to ensure that children stay engaged and complete the entire experiment.

Participants

We preregistered a planned sample of 60 children ages 4-8. At the time of writing, 15 participant responses have been partially transcribed and analyzed.

Procedure

Experiment 3.2 is being conducted asynchronously. All information and instructions are prerecorded, and a slide-style Qualtrics survey is hosted on the Princeton and NYU Discoveries in Action Lab (PANDA). Children are introduced to a fictional younger child, "Sam," in the same manner as in Experiment 4.1. However, instead of providing *ball* and *piano* as examples of words Sam knows and does not know (respectively), children are told that Sam knows *duck* and *penguin*. As such, for *duck* and *penguin* trials, children do not need to make real-time inferences about Sam's knowledge.

To familiarize children with "communicating" with Sam, children were played an audio of "Sam" saying hello ("Hi, I'm Sam"), and encouraged to greet Sam. Children are told when to speak, and how to press a button to relay the message to Sam. During the task, children are reminded to press the button whenever they have finished speaking.

After greeting Sam, children were told that they need to help Sam place stickers into a sticker book. Participants are told that they have the list with the correct orders, and that they must guide Sam to select the correct stickers. Prior to the task, and at the beginning of each trial, children are told that Sam cannot see the images on children's screens, and that they must verbally tell Sam what sticker to find.

Inference and no-inference trials. On all trials expect *duck* and *penguin*, children would be required to make real-time inferences about Sam's knowledge of that particular animal. For *duck* and *penguin* trials, children are reminded at those specific trials about Sam's knowledge or ignorance. The *no-interference* trials mimic the procedure of Baer and Friedman's (2018) study, where children are explicitly told their partner's knowledge state. On all trials, children hear "This is the sticker that Sam needs to find next. Remember, Sam can't see what's on your screen, so you need to use your words to help Sam find that sticker. Ok, here's Sam. Now it's your turn to talk to Sam. Remember to press the button to send the message to Sam when you're done talking!" On *no-inference* trials, children also hear "Look! This is a [penguin]. Sam [doesn't know] that this animal is called a [penguin]."

AoA trials. After the 10 communication trials, children were asked to give judgments about Sam's word knowledge of the same 10 animals. To shorten the length of the experiment, children were only asked to give binary (*yes* or *no*) judgments, rather than the two-step scale used in Experiment 4.1. The AoA trials served two purposes: 1) to investigate any adaptation or

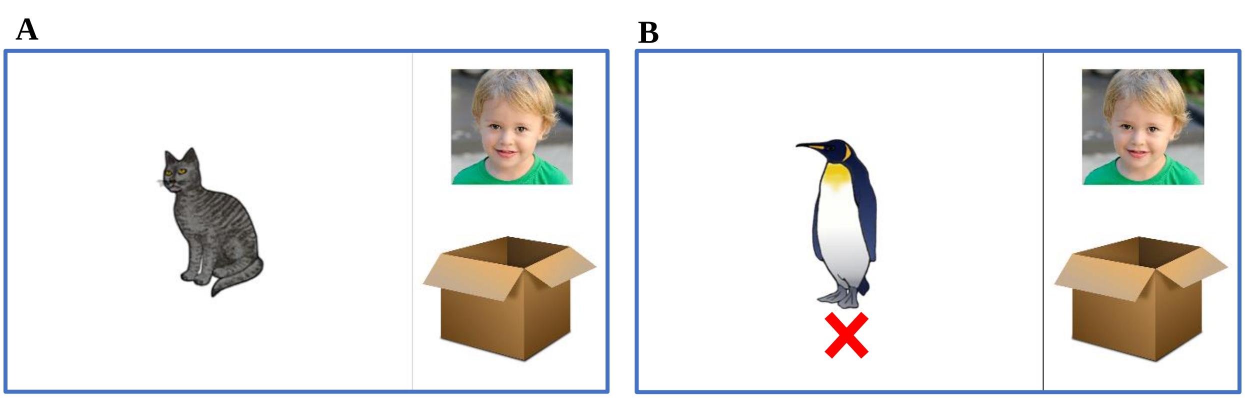

Figure 3.6 *Example trials. (A) On inference trials, children are not told whether Sam knows the animal. (B) On no-inference trials, children are reminded whether Sam knows or does not know the animal.*

modification of speech based on children's specific estimates of Sam's knowledge; 2) as a manipulation check to ensure that children remembered that Sam knows *duck* but not *penguin*.

3.2.2. Preliminary Results and Discussion

Data collection for Experiment 3.2 is still ongoing, and I will briefly discuss preliminary findings here. As a first step in uncovering whether children engage in communicative adaptation, we transcribed participants' speech during the *duck* and *penguin* trials only. These trials most closely mirror prior work by Baer and Friedman (2018), and do not require children to make real-time inferences. Thus, communicative adaptation may be more likely, as children are not burdened with the additional cognitive load of making inferences.

Contrary to our expectations, children of all ages in this partial sample do not appear to modify their speech regardless of Sam's knowledge. That is, even when explicitly reminded that Sam does not know what a *penguin* is, children do not adjust their descriptions in response. While there are variations between children in how they communicate with Sam, there is little within-subject difference between *duck* and *penguin* trials. Currently, due to the exploratory nature of these preliminary analyses, adaptation is coded based on researcher judgment.

If our preliminary findings persist, what might they mean? Since children are predicted to adapt *more* on *no-inference* trials, it is likely that children in our sample also do not adapt on *inference* trials. One possibility is that, despite a sophisticated ability to make fine-grained estimates about another person's vocabulary knowledge, children lack the ability to leverage that in real time during communicative settings. Given that children do not appear to adapt even when their partner's knowledge is provided, it may be that children are unable to generate alternative descriptions of the animals. This account would be in line with the findings from Experiment 2.3: children overwhelmingly used the same label when asked to uniquely identify familiar objects, regardless of context (e.g., saying “dog” when there is a dog and a rabbit on the screen, as well as when there are two dogs present). Children may have trouble hindering the most salient label for a given animal and fail to adapt to a less knowledgeable partner.

A second interesting possibility is that *live* interaction facilitates communicative adaptation. The asynchronous nature of Experiment 3.2 means that participants do not receive any feedback from “Sam,” which likely departs from children’s daily interactions with others. Further, participants are not told whether Sam has correctly selected any of the stickers. Future studies should utilize live interactions, either with genuine participants, confederates, or even prerecorded responses from an actor child, to more closely mimic the communicative instances that children engage in every day.

Chapter 4. What facilitates communicative perspective-taking? An online computation account of communication

Successful communication requires that conversational partners understand one another. What allows conversations to be successful? Various models of communication have been proposed to account for the factors that drive communicative success (and errors). These models are formulated and tested using data from adults, but adults are not the only people who communicate—even young children engage in conversation with those around them. While a wide range of studies in communication has been done with both adults and children, theoretical work has thus far focused on adults. Developmental work can be critical not only for understanding how children develop communicative abilities, but also in informing our theories about communication. A comprehensive framework of conversation should explain and make plausible predictions not only about communication in adults, but in children as well. In this paper, I discuss three models of communication, and evaluate them against data from adult and developmental work. Further, I propose an *online computation account* of communication. Under this model, information required for perspective-taking is either computer *in the moment*, or *pre-computed*. As such, perspective-taking is facilitated by the amount of information that is already stored and easily retrieved (i.e., *pre-computed*), and potentially hindered by the amount of information that is new and must be considered during online processing (i.e., *in the moment*).

Three Models of Communication

Conversation involves at least one speaker and one listener, who each have communicative pressures that may dictate how they produce and/or interpret information. Further, each conversational partner has their own (literal and figurative) perspective. Successful communication thus relies on perspective-taking—speakers and listeners account for each other's pressures and knowledge in order to effectively convey their messages. How does perspective-taking occur in conversation? One possibility is that perspective-taking occurs late in processing. This account, proposed by Keysar (2007), posits that speakers default to considering their own perspectives, and only incorporate partner perspectives after monitoring. Hanna, Tanenhaus, and Trueswell (2003) offer a second account, arguing that perspective-taking is probabilistic. Listeners consider multiple sources of information during communication, adjusting their weighting to ultimately determine the most likely meaning of a speaker's utterance. Finally, a third model proposed by Hawkins, Gweon, and Goodman (2021) begins to unify these accounts by considering how both speaker and listener pressures and perspectives influence communication. This resource-rational model posits that conversational partners divide the "labor" of communication amongst themselves. Speakers and listeners each consider their own and their partner's pressures and perspectives, and determine the most efficient weighting of each perspective.

How well do current theories explain data from developmental studies? The empirical basis for current theories of communication lies largely in adult data. To my knowledge, only one developmental study directly engages with existent models of communication (Nadig & Sedivy, 2002). A developmental perspective is crucial not only because it can inform our understanding of how children learn to communicate effectively with others, but also because it provides a "special" case for which to test our theories of communication. Children regularly interact with other conversational partners, but their communicative contexts differ from those of adults. First, children do not yet have the cognitive and linguistic capabilities of more fluent speakers. While both production and comprehension are likely influenced by children's own developmental level, children may be especially constrained when it comes to production. Second, children often interact with others that are significantly more knowledgeable and communicatively competent than they are, creating an interactive context that is highly asymmetric.

Developmental Findings

Many developmental studies have asked *when* children acquire certain communicative skills, but a more important question may be *how* children's developmental trajectories can lend insight into current theories of communication. Specifically, how might developmental work aid our understanding of how and when perspective-taking occurs in language processing? One perplexing feature of current developmental data is the apparent contradictions in results—while some studies show that children struggle massively in communicative contexts, others have shown that they readily account for other's perspectives in communication. How do existent theories of communication explain these discrepancies across studies, and what are the factors driving children's success and difficulties in communication?

Children appear to struggle with a variety of communication tasks even through middle childhood, possibility due to an inability to coordinate differing perspectives between themselves and their conversational partners (e.g., Krauss & Glucksberg, 1977). In a series of studies, Krauss & Glucksberg (1977) tested children's ability to communicate with peers in director-matcher tasks, where participants took turns guiding their partner to select an abstract object while not being able to see one another. The authors found not only that preschool-age children performed poorly on director-matcher tasks, but that even older children made significant and consistent errors. One possibility for why children struggle in these tasks is that they are incapable of identifying sufficiently informative expressions (Asher & Oden, 1976; Robinson & Robinson, 1977). Alternatively, communicative failure may arise when children do not offer sufficient feedback to their partner or ask for clarification when necessary (Anderson, Clark, & Mullin, 1994). Yet another possibility is that the difficulty of describing abstract shapes, which are often used in director-matcher tasks, interferes with children's ability to communicate effectively.

On the other hand, various studies have shown that children can and do communicate effectively with others by accounting for shared knowledge and partner perspective in language production and comprehension (Branigan, Bell, & McLean, 2016; Nadig & Sedivy, 2002). Even preschool children keep track of difference in visual perspectives between themselves and their conversational partners, such that they reliably fixate on a correct referent when a distractor is present but not visible to the partner, and are less likely to use disambiguating adjectives to describe an object when a distractor is hidden from their partners' view (Nadig & Sedivy, 2002). This study suggests that children, at least when interacting with an adult, can reliably account for others' perspectives in communication. Branigan et al. (2016) examined children's ability to coordinate with peers in a director-matcher task, and found that by 8-10 years of age, children form referential pacts with others, a well-studied phenomenon in adults (see Clark & Wilkes-Gibbs, 1986). Furthermore, children's referential pacts are partner-specific, such that children adjust referential expressions when a new conversational partner is introduced. What might be driving children's success in these studies? One factor may be theory of mind, or the ability to reason about others' mental states, perspectives, and beliefs. One study found that children's performance on theory of mind tasks predicted communicative efficacy in a director-match tasks, suggesting that these abilities may co-develop in childhood (Sidera, Perpiña, Serrano, & Rostan, 2016). Executive functioning, specifically inhibitory control, also predicts children's perspective-taking abilities in communication, likely due to the need to inhibit one's egocentric perspective when considering a partners' referential expressions (Nilsen & Graham, 2009).

How do we explain the differences in empirical data? While factors such as age of participants and task difficulty likely contribute to some variability in children's communicative success, little work has addressed the potential theoretical implications of current findings. What, if any, current theories can account for children's (and adults') selective success in perspective-taking? In the following sections, I consider how current frameworks of communication may explain developmental and adult data, with particular focus on how these frameworks account for the apparently contradictory results from developmental work.

Account 1: Two-Step Models of Communication

One account of communication posits that language production occurs in two stages. Specifically, this theory posits that the initial stage of language processing in egocentric, and other perspectives are incorporated after monitoring (Horton & Keysar, 1996; Keysar, Barr, Balin, & Brauner, 2000). Under this model, speakers appeal first to egocentric perspectives, which may arguably be sufficient in many situations—we often share our partners' perspectives, and thus no additional information about others' beliefs may be necessary for effective communication. This account is appealing due to its simplicity, as well as computational efficiency. If relying on egocentric perspectives leads to communicative success most of the time, it would be unnecessary and inefficient for others' perspectives to be taken into consideration.

Studies of perspective-taking in both children and adult indicate that egocentrism can interfere with and hinder communication. Numerous studies show that children struggle to adapt to partners in communicative tasks (e.g., Krauss & Glucksberg, 1977), and some of these difficulties may arise from an inability to inhibit their egocentric perspectives (Nilsen & Graham, 2009). Even more compelling evidence comes from work with adults. In a matcher-director task where most items were in shared visual ground (i.e., both partners could see these item), but some were in privileged ground (i.e., only one partner could see these items), adults were as likely to look at an object in privileged ground than in shared ground when both matched the description given by the director (Keysar et al., 2000). That is, even though an object was hidden from the speaker, and thus could not plausibly be the referent, listeners still considered the hidden object as the possible referent. This surprising result indicates that egocentric perspectives are not ignored and may interfere with language comprehension. Similar errors, or failure to inhibit one's own perspective, are found in language production. In one study, speakers engaged in egocentric processing when under time pressure, producing over-informative expressions when describing objects to their partner (Horton & Keysar, 1996). Crucially, speakers were less likely to be over-informative when there was no time pressure, suggesting that the pressure may have disrupted their language processing and led them to rely on the egocentric "default."

The two-step account provides an intuitive way of understanding communicative errors and failures. In the existence of factors such as cognitive load or distractions, language processing is disrupted and does not progress to the perspective-taking stage. This account appears to offer a satisfying and plausible explanation for children's failure in various communication tasks. Given that children's cognitive capacities may be more limited than adults', any communicative task may be more difficult for children to navigate. Thus, the failures of children to coordinate successfully with a partner even into middle childhood (e.g., Krauss & Glucksberg, 1977) might be because children fail to reach the effortful stage of perspective-taking, instead reasoning solely from their egocentric perspective. Communicative performance does improve with age, improved executive functioning, and more sophisticated theory of mind, all of which are plausible under an "egocentric-first" account. As children develop, their improving cognitive capabilities means that tasks become less difficult, and while they may still suffer from the egocentric "default" from time to time, they are nonetheless more reliable communicators than younger children.

How does this two-step model account for evidence showing that young children *do* rapidly take others' perspectives into account? One way to explain children's success would be that certain communicative tasks are easier. For example, describing common household objects is likely easier than describing abstract shapes. Thus, the cognitive load of the task at hand may not impede children's perspective-taking abilities. However, simply claiming that all of these experiments are necessarily "easier" is not a particularly satisfying explanation. Furthermore, Nadig and Sedivy's (2002) study offers compelling evidence that perspective-taking occurs early on in processing, even for 5- and 6-year-old children, which does not fully lend support to the two-step model's argument that incorporate others' perspectives is necessarily effortful and late-occurring.

Account 2: Probabilistic Models of Communication

A more nuanced theory proposed by Hanna, Tanenhaus, and Trueswell (2003) posits that while egocentric perspectives are not completely ignored, other perspectives are also taken into account at the earliest moments of language processing. In their study, Hanna et al. (2003) investigated the time course of perspective-taking using eye tracking. In two experiments, participants were paired with a confederate speaker to play a communication game in which the speaker instructed the participant to pick up an object (e.g., a martini glass). Importantly, participants were aware that certain objects were in privileged ground, i.e., that only the participants saw or knew about those objects. For example, in one condition, the experimenter mislabeled an object that only the participant had visual access to. As such, participants were aware that the speaker had an incorrect representation of that object. The crucial analysis in this set of experiments was to compare eye tracking data between privileged ground and shared ground conditions. Privileged ground trials involved objects that only the participant could see, whereas shared ground trials involved objects visible to both partners. Hanna et al. (2003) found that while participants were slower to fixate on the target object in the privileged ground condition than in the shared ground condition, they were faster in the privileged ground condition compared to an ambiguous trial. Crucially, participants fixated more on the correct target than the object in privileged ground, even when the object in privileged ground was a better match for the speaker's description. That is, eye tracking data revealed that adults rapidly account for other's perspectives in language processing, even though egocentric perspectives are not completely ignored.

This view of perspective-taking in communication might reconcile the seemingly contradictory results in the developmental literature. Under this account, incorporating a partner's perspective is not an "all-or-nothing" process. Rather, information is processed in real-time, and each partner's perspective is weighed during language processing. Importantly, while the account claims that multiple sources of information are available and *can* be taken into account early on in processing, it does not claim that individuals necessarily adapt to others' perspectives in communication. That is, egocentric errors can and do occur, but the reason it occurs is not due to disruption in language processing.

This framework of incremental processing may help explain the wide range of empirical results in the field. In reviewing and re-analyzing multiple eye-tracking studies, Brennan and Hanna (2009) found varying time-courses of perspective-taking across studies. While some studies showed perspective-taking at very early stages of language processing (180-300ms; Brown-Schmidt, 2008; Brown-Schmidt, 2009), others showed slower time-courses (e.g., 900-1200ms; Metzing & Brennan, 2003). This wide range of response times in perspective-taking appears to lend support for an incremental processing account, rather than an "all-or-nothing" account. If there is an egocentric "default" that individuals fall back on when perspective-taking is disrupted, one might expect a more uniform time-course across situations for when other perspectives are taken into account. On the other hand, the incremental processing account would predict variability in time-courses of perspective-taking, as the time needed to incorporate and weigh different sources of information varies based on the communicative contexts that individuals are in.

A probabilistic processing account may also be more suitable for understanding the mixed results in studies with children. On the one hand, children as young as 5 may be able to account for differences in visual perspective in both language comprehension and production, suggesting an early-emerging and sophisticated ability to keep track of multiple perspectives (Nadig & Sedivy, 2002). On the other hand, some evidence suggests that even 9th-graders struggle to coordinate with peers on matcher-director tasks involving novel shapes (Krauss & Glucksberg, 1977). The probabilistic processing account predicts that multiple perspectives are available to conversational partners, but whether they are fully incorporated or ultimately used depends on the weights of each perspective. Under this probabilistic model, children's errors arise not from complete disruption of the language processing system, but due to factors influencing the computation of which perspective to weigh more heavily in any given context. For example, some paradigms may bias children to weigh their own perspectives more heavily, and potentially lead to more egocentric errors.

Account 3: Resource-Rational Models of Communication

Thus far, the models discussed have focused on communicative instances, often at the level of disambiguating meaning in a sentence. However, conversations are much more than disjointed sentences pieced together. How do conversational partners dynamically adapt to one another throughout the course of an interaction? A recent framework proposed by Hawkins, Gweon, and

Goodman (2021) posits that conversational partners are resource-rational—speakers and listeners weigh the relative benefits and costs of adapting to their partners, in relation to what they expect their partners to do. As such, conversational adaptation can be seen as a *division of labor* between speakers and listeners: when a speaker expects a listener to weigh their own perspective heavily, speakers may then contribute more effort to adapt to this listener. In two experiments, Hawkins et al. (2021) found that conversational partners dynamically adjust the degree of perspective-taking they engage in, based on their partner's behavior. This view of communicative adaptation provides a holistic view of communication—speakers and listeners each consider the costs and benefits of perspective-taking, ultimately to minimize effort while maximizing communicative success. Under this account, communicative errors arise when speakers and listeners are misaligned in how much effort they expect the other to contribute. Crucially, speakers and listeners can account for these misalignments and adjust over time to reduce errors.

A similar model of communication was put forth by Ryskin, Stevenson, & Heller (2020). The authors investigated how speakers and listeners weight own and partner perspectives, and found that speakers weighted their own perspective more, whereas listeners weighted partner perspectives more. This asymmetry may suggest that speaking is more cognitively taxing than listening, and thus warrants accommodations from the listener. In line with Hawkins et al.'s (2021) framework, Ryskin et al.'s (2020) study suggests that speaker and listener perspectives are considered by both parties, but the specific weighting may depend on which party is "carrying the burden" of communication. In director-matcher paradigms where speakers are required to consider multiple objects in an array, while calculating (potentially) mismatched perspectives, listeners (or matchers) may be expected to carry more of the communicative load by adapting to the speaker.

A key strength of the "division of labor" framework is that it considers both the listener and speaker and allows for weights to shift dynamically over the course of communication. Conversation is inherently cooperative, and a model that accounts for both listener and speaker behavior may be more capable of capturing observed empirical data than a model that only focuses on one conversational partner. This line of work provides a promising avenue to understanding children's developing communicative abilities. A "division of labor" model of communication may help explain why children succeed in communicating with adults, especially familiar adults, but not with peers. Taking parent-child pairs as an example, parents may weight the child's perspective more heavily to scaffold successful communication. Children, who interact with their parents constantly, may have an expectation that they will be understood. Parents and children may have success in communication because they are aligned in their expectations with regards to the division of labor. More generally, adults may be able to communicate effectively with children because they are broadly aware of children's cognitive and linguistic abilities, and assume a greater role in supporting communicative success. Some support for this account comes from Glucksberg, Krauss, and Weisberg (1966), who showed that preschool children succeeded in director-matcher tasks when adult experimenters used the labels children generated for the novel referents. Here, the "labor" of communication falls mostly on the adult, who intentionally reused children's (idiosyncratic) labels in order to support communicative success. Under this framework, children's failure to communicate with peers about novel referents could be due to misaligned expectations of their own and their partner's efforts.

Limitations of Current Models of Communication

Thus far, I have discussed three theoretical frameworks for understanding perspective-taking in communication, and how each of these theories may explain data from both adult and child studies. A two-step model offers intuitive explanations as to why errors may arise in communication, but this model provides little nuance. An incremental processing model may better capture the wide variance in the time-courses of perspective-taking in communication but lacks predictive precision. Finally, a resource-rational model provides a framework for understanding how speakers and listeners coordinate multiple perspectives, and is a promising avenue to understanding the dynamic nature of human communication. However, an outstanding question remains: *What* drives the incorporation (or failed incorporation) of multiple perspectives?

The two-step model makes binary predictions—either individuals will take common ground into consideration, or not. Thus, this model cannot make nuanced predictions, and may rely on a priori assumptions to make certain predictions. For example, a task may appear more difficult and cognitively tasking than it actually is. A researcher reasoning from a two-step framework may thus predict a high rate of egocentric errors. If the data turn out to show that participants rarely made egocentric errors, the researcher could simply conclude that the task was not actually difficult. Therefore, with the exception of experiments designed to specifically test this theory, the two-step account may not offer insight into many of the other studies of communication.

While more nuanced, and better able to capture the existing data, the incremental processing theory also lacks predictive precision. The model appears to better explain the variance in results across studies, and may be a better description of the processes underlying perspective-taking in communication. However, due to its descriptive nature, it is unclear how this model should make predictions for new studies. Hanna et al.'s (2003) experiments cleverly showed that multiple sources of information are available to individuals even early on in processing, but their theory does not make claims about which sources of information should be weighted more heavily, and whether all sources of information are available on the same timescale.

Finally, the resource-rational model is perhaps the most comprehensive theory thus far. It explicitly considers both speaker and listener pressures, and view these pressures as dynamic. This model of communication suggests that both speakers and listeners weight each other's perspectives during online communication. Importantly, speakers and listeners rationally adjust how much of the perspective-taking "labor" they take on while maintaining successful communication. This model could predict the variance in communicative success (seen in both adults and children) but is largely agnostic to the mechanism that leads to certain divisions of labor.

Importantly, none of the discussed here have directly engaged with developmental questions. Above, I have considered how each model may account for existent developmental data, but very few studies approach children's communication from a theoretical perspective. To that end, existent theories are also largely agnostic regarding predictions about children. Further, the three discussed models focus on *when* communication is successful (or not) but not on *what* drives success. Frameworks that rely on situational factors may be especially difficult to test on children due to the challenge of identifying the underlying reason for children's communicative struggles: it is not always clear whether children simply lack the cognitive or linguistic capabilities to communicate, or if there are situational factors (similar to those facing adults) that are driving their perspective-taking.

An Online Computation Account of Communication

I propose a potentially unifying account of communication that focuses on *in-the-moment computation* as a key factor guiding perspective-taking and adaptation in conversation. In the present framework, *in-the-moment* computation refers to any information that speakers and listeners must incorporate *during* the conversation to communicate effectively with each other. For example, for an individual to refer to an object correctly for their listener, their in-the-moment computation may include the total number of possible referents in their visual field. Additionally, where visual perspectives differ, the speaker must also compute what their partner sees. In line with some of the above theories, my framework predicts that the complexity of the environment (e.g., how many possible referents there are), as well as whether interlocuters share the same perspective, should influence computational load and thus perspective-taking behavior.

A crucial component of the proposed account is that in-the-moment computational load does not necessarily increase with sheer amount of information. In some cases, complex and rich information about a partner can be *pre-computed*, thus allowing for rapid perspective-taking. While online processing pressures exist in all communicative contexts, much of the computational load can be relieved if a considerable amount of relevant information is already computed. One way that information might be precomputed is through prior interactions. If conversational partners have already established shared ground, this shared information may be immediately available because it has already been computed before. For example, when a speaker engages with a listener who they have not met before, they may need to estimate the listener's knowledge *in the moment* to convey a message successfully. However, when the same speaker engages with a familiar listener, the speaker may simply retrieve *pre-computed* information about this listener's knowledge. Therefore, in many cases, perspective-taking may appear to occur rapidly and even automatically.

A similar model of online perspective-taking has been proposed by Galati and Brennan (2006; see also Brennan & Hanna, 2009). The authors posit that some communicative contexts are computationally easier to process, and thus allows speakers and listeners to spontaneously account for each other's perspectives. In these contexts, Galati and Brennan (2006) propose that interlocuters may build "one-bit models" of their partner, leveraging one key piece of information that allows them to readily compute another partner's perspective. For example, whether a partner shares visual access to the same objects (Nadig & Sedivy, 2002), or whether a partner is hearing a story for the first time (Galati & Brennan, 2006). In these situations, one (often binary) "bit" of information is particularly relevant when considering the partner's perspective, which facilitates quick computations of and immediate adaptations to partner perspectives.

The online computation account is arguably a more generalizable version of one-bit models, in that it offers explanations for a wide range of empirical results. A key difference between the one-bit models and the online computation account is that one-bit models focus on amount of information rather than amount of computation. While conceptualizing information as "bits" is a plausible explanation for how listeners and speakers seemingly adapt spontaneously to one another in conversation, this framework relies on the assumption that all types of information are equally effortful computationally. That is, as long as the relevant information is one "bit," it should not matter whether that information itself is easy or difficult to compute. Consider two types of information: the age of a partner and the visual perspective of a partner. While each of these can be construed as one "bit," it seems unlikely that incorporating another person's visual perspective (possibly of multiple objects) is as easy as considering whether they are a child or an adult. Thus, one-bit models may offer a broad explanation for why perspective-taking is easier in some situations than others, but fails to fully address the gradient time-courses found across multiple studies. On the other hand, focusing on the *amount* of computation required (regardless of amount of information) may be more appropriate for both understanding the existing data and making plausible predictions for communicative behavior. As such, one-bit models could be considered one instantiation of the online computation account—when only one piece of information needs to be computed in-the-moment, the low computational demand facilitates rapid perspective-taking.

The amount of computation required during a communicative interaction is no doubt influenced by the amount of information that an individual needs to consider. However, it need not necessarily be the case that more complex information always requires more computation in the moment. Under certain conditions, such as with abundant prior experience with the same communicative partner, it may be possible to quickly compute and account for the partner's perspective. As such, some communicative situations may require individuals to consider a great amount of information, but computational effort can still be low if these pieces of information have already been encountered before. One line of work that fits well into this framework is research on parent-child communication. Evidence shows that parents are highly sensitive to their own children's specific vocabulary knowledge and can leverage their complex models of their children's knowledge in spontaneous communication (Masur, 1997; Leung et al., 2021). These findings suggest that some amount of information in communicative interactions can be pre-computed, and thus accounting for a partner's different perspective (or knowledge) can be spontaneous and relatively effortless.

The online computation framework is similar to and compatible with the incremental processing account in some ways, in that perspective taking is not an "all-or-nothing" process, and that information required for perspective taking is available to conversational partners immediately (and even before an interaction begins). By introducing *in-the-moment* computational effort as a parameter that constrains or facilitates perspective taking, our framework seeks to further understand how and when perspectives are taken into account in communication. More importantly, the goal of the present framework is not only to capture and explain existing data, but to make clear and specific predictions about how adults and children account for others' perspectives in conversation. A key contribution of this framework is that it extends to more naturalistic situations of communication. Describing objects that are partially obscured from a partner makes up very little of our daily interactions with others—rather, the common ground that we must consider in conversation often includes other aspects of our partners' knowledge.

In the following sections, I will show how the proposed online computation account offers explanations for some of the communicative phenomena described above. I will show that the online computation account can be applicable to data from both the adult and developmental literature. Additionally, I will propose predictions that this account makes regarding perspective-taking in online communication.

How does the online computation account explain current adult data?

The adult literature on communicative perspective-taking has yielded a wide range of results, which has led to varying conclusions about the nature and time-course of perspective-taking. One merit of the online computation account is its potential to explain and provide insight into why different experiments should yield seemingly contrary results. When is perspective-taking easy and seemingly automatic, and when is it effortful and slow? The online computation account posits that in-the-moment computational load drives the ease (and difficulty) of online perspective-taking.

One line of work shows that even adults "fail" to rapidly account for another's visual perspective in referential communication tasks, and these results have been used to support arguments for an egocentric "default" in language processing and production (e.g., Horton & Keysar, 1996; Keysar et al., 2000). However, these results are also in line with an online computation account. In Horton and Keysar's (1996) study, the main manipulation was time pressure—the authors posited that if egocentric processing precedes perspective-taking, then time constraints would disrupt perspective-taking abilities. Adults did appear to produce more egocentric references when under time pressure, but this outcome need not have been due to an egocentric default. A theory positing an encapsulated egocentric stage in language processing should not only predict that earlier processing is more likely to be egocentric, but that there should be a set time "limit" for when an individual can and cannot make use of other perspectives. As discussed above, the varying eye-tracking data from multiple studies does not suggest a consistent time-course for when perspective-taking can occur. It is thus possible that the observed results were due to disruption in computation—the time pressure demands may have hindered individuals' abilities to complete the necessary computations and weighting of perspectives.

The online computation account also predicts that there should be changes in perspective-taking behavior over time, depending on the communicative and referential context. Some paradigms discussed above required a new computation on each trial (e.g., Horton & Keysar, 1996). If the possible set of referents changes from moment-to-moment, individuals will need to re-compute their own perspective and their partner's perspectives in-the-moment. Therefore, while practice may lead to slightly better performance due to familiarity with the task, the computational load does not change over time. On the other hand, some director-matcher paradigms do not require completely new computations on each trial (e.g., Keysar et al., 2000). Our account predicts that when computations must be done from scratch, time spent with the same partner should not significantly decreased the computational load, and thus perspective-taking behaviors should stay relatively the same over time. However, when listeners and speakers engage in conversation about the same visual scene for an extended period of time, their perspectives (whether shared or not) do not need to be re-computed on each trial. Our theory predicts that over time, individuals will exhibit fewer errors in perspective-taking because partner perspectives become established in common ground and are pre-computed, rather than computed from scratch during each conversational turn.

How well does the prediction that multiple interactions with the same partner will reduce errors map onto empirical data? Using data from prior studies, we can explore whether 1) maintaining the same context (e.g., visual array of objects) and 2) interacting with the same individual over time reduce perspective-taking errors. Our theory predicts that communicative contexts requiring new computations may hinder perspective-taking due to high computational load. In one study, Keysar et al. (2000) tested perspective-taking behaviors when an array of objects changed from trial-to-trial, as well as when the array remained the same. When directly comparing the two experiments, there was no evidence that maintaining the same array reduced the overall error rates. However, since trial level data are not available, it is unclear whether error rates changed over time. While the authors did not find differences in overall error rates between experiments with a changing or static display, it is possible that participants in the static context improved their perspective-taking over time. Alternatively, it could be that the complexity of the visual array required even more trials before individuals could fully incorporate their partner's perspective into their precomputation. In a recent study, Hawkins et al. (2021) found evidence for multiple interactions leading to a decrease in perspective-taking errors. Using a modified paradigm from Keysar et al. (2003), Hawkins et al. (2021) replicated the finding that individuals make errors in perspective-taking when a distractor is present in privileged ground. Crucially, Hawkins et al. (2021) found that error rates decreased as the number of interactions with the same partner increased. That is, over time, individuals adopted more of their partner's perspective relative to their own. The online computation account proposes that this pattern of results arises from a reduction in computational load—repeated interactions with the same partner in the same communicative contexts allows for partner perspectives to become precomputed, thus lessening in-the-moment computation.

The online computation account's predictions also extend to naturalistic conversations. Our daily interactions often involve people who we have encountered before, and a general theory of communication should extend to these situations. While laboratory studies often use confederates, or pair participants with strangers, parent-child interaction studies may provide a special case study for understanding perspective-taking and adaptation in natural conversation. Parents and young children differ greatly in their knowledge and linguistic abilities, thus adaptation to one another could be particularly difficult. At the same time, parents and children communicate with each other constantly. The online computation account predicts that repeated interactions with the same partner can facilitate precomputation, as knowledge about the partner increases. Parents should thus be well-adapted to their children and able to rapidly take their perspective in communication, despite the great distance between the two parties' abilities and knowledge. Indeed, both observational and experimental studies show that parents adjust their speech according to their beliefs about their children's vocabulary (Masur, 1997; Leung et al., 2021). While response times were not measured in either of these studies, parents appeared to spontaneously produce sentence structures or content that depended on what they believed their children knew. In Leung et al. (2021), the parent-child game had a clear communicative goal—parents were instructed to help their children find a target animal amongst a set of three. While time pressure was not a variable, experimenters noticed that parents likely experienced some pressure to communicate quickly, due to young children's eagerness to engage with the iPads used in the game. Thus, while the effect of time pressure was not explicitly measured, Leung et al.'s

(2021) study shows that parents rapidly account for their children's knowledge, and tailor their speech accordingly.

How does the online computation account explain current developmental data?

Children's ability to communicate with others has been studied using a wide range of paradigms. While many of these studies aimed to understand when certain competencies emerge, few studies focus on the theoretical implications of children's successes and failures in various communicative tasks. Developmental research on communication is not only important for our understanding of children's trajectories, but can also inform our theories for communication. The online computation account seeks to capture and make predictions not only about how adults communicate, but also how children do so (or fail to do so). This theoretical framework can also bring together multiple lines of work and shed light onto the broader picture of children's communicative development. In the following section, I will show how the online computation account might explain seemingly contradictory findings in the developmental literature.

While a classic set of studies by Krauss and Glucksberg (1977) showed that even teenage children struggle to coordinate with their partner on a communicative task, a number of studies have shown that young children are able to adapt to their conversational partner's perspective in both comprehension and production (e.g., Nadig & Sedivy, 2002; Branigan, Bell, & McLean, 2016). One factor that could be driving the differences in results across studies is the stimuli. Whereas children in Krauss and Glucksberg's (1977) study were asked to talk about abstract shapes to one another, participants in Nadig and Sedivy's (2002) experiment were communicating about familiar objects. The complex stimuli in Krauss and Glucksberg's (1977) may have increased children's cognitive and computational load during the study, hence impacting communicative performance. Additionally, differences in experimental paradigm may contribute to in-the-moment computational pressure. In eye-tracking tasks, participants can see which objects are occluded from their partners during online production and comprehension. While this may still require some computation during processing, information about what is and is not occluded is immediately available to participants. On the other hand, matcher-director tasks similar to Krauss and Glucksberg's (1977) require participants to be separated by a divider. The inability to see a conversational partner may lead to increased difficulty in simulating the partner's perspective and could impact communicative perspective-taking: 1) speakers may simply forget that the partner's perspective is different, and thus neglect to engage in perspective-taking, or 2) the computational load is increased when conversational partners are not visible to one another, hence hindering perspective-taking behavior. Both of these effects would be particularly pronounced in children, who may already be struggling with the cognitive and computational demands of the task.

One interesting finding from multiple studies is that children's performance on communicative tasks improves when they receive feedback (Branigan et al., 2016; Krauss & Glucksberg, 1977; Matthews et al., 2007). In matcher-director tasks that involve multiple rounds of communicating about the same novel referents, children's accuracy improved over time when given feedback from the experimenter (Branigan et al., 2016; Krauss & Glucksberg, 1977). This pattern of results is in line with adult data (e.g., Hawkins et al., 2021), and is consistent with the online computation account's prediction that accuracy should increase over multiple interactions with the same partner. Since both Branigan et al. (2016) and Krauss and Glucksberg's (1977) studies used novel referents and required conversational partners to be obscured from one another, adult feedback about performance may have been necessary to help children communicate successfully with one another. Experimenter feedback may have reduced the in-the-moment computational demands during the task, and provided children with the key information (i.e., whether their communication on a prior round was successful) they need to adapt to their conversational partners.

Predictions Made by the Online Computation Account

While potentially compatible with the online computation account, the developmental studies discussed above do not allow for a clear distinction between computational and cognitive load. That is, much of the developmental findings could be explained by an alternative account that cognitive load drives communicative perspective-taking, where perspective-taking occurs more quickly when cognitive pressures are lower. While online computational load is likely tied to general cognitive demands, the online computation account predicts that certain conditions would allow for rapid perspective-taking, irrespective of cognitive load. Specifically, when some sources of information are pre-computed, they can be leveraged easily during online communication, thus translating into rapid perspective-taking. For example, the online computation account would predict that children interacting with a stranger would struggle more to engage in perspective-taking, when compared to children interacting with a familiar partner. Evidence showing that children's error rates gradually decrease over time when interacting with the same partners (and increase when a new partner is introduced) may lend some support to this argument (Branigan et al., 2016; Krauss & Glucksberg, 1977), but further research should specifically probe the effect of prior interaction on perspective-taking behavior in online conversation.

The online computation account also predicts, similar to Galati and Brennan's (2006) one-bit models, that children can rapidly adapt to their conversational partners when the necessary computations are simple. Do children adjust their speech according to their conversational partners? One study by Shatz and Gelman (1973) suggests that children do so at a surprisingly young age.

By age 4, children describing a toy to an adult used longer sentences than when describing the same toy to a 2-year-old. This result most directly lends support to the idea of one-bit models: in the study, the relevant piece of information that the 4-year-old speakers needed to compute was whether their listener was an adult or a younger child. Children could then leverage their general understanding of adults' and young children's abilities to communicate effectively. As discussed in previous sections, the online computation account can be seen as a generalized version of one-bit models, and thus may also glean support from Shatz and Gelman's (1973) findings. It is important to note that the children in Shatz and Gelman's (1973) study interacted with unfamiliar 2-year-olds, and not their own siblings. To test the plausibility of the online computation account, future work should compare children's interactions with familiar versus unfamiliar younger children. If children do form rich models of their social partners, and can leverage those models in online communication, perspective-taking behavior should arise more quickly when children are interacting with their own siblings than with unfamiliar children, even if adaptation to the listener is seen in both cases.

The online computation account differs from many current frameworks of communicative perspective-taking in that it gives precise, gradient predictions for behavior. Specifically, the online computation account suggests that ease of perspective-taking relies on how much information can be pre-computed. As such, amount of prior interaction with a partner is a key factor influencing perspective-taking behavior. Further, the model predicts similar behavior patterns in children and adults. While there may be differences in cognitive and linguistic abilities between adults and children, the overarching driver of perspective-taking is online computational load. Take a parent-child pair talking to each other about abstract shapes: while children may struggle to describe the shapes, they may succeed in communicating with their parent because of their wealth of prior interactions. Children and parents know the common knowledge that they share, as well as the idiosyncratic behaviors and perspectives each other has, and can thus rapidly take the other's perspective. A parent may quickly identify an abstract shape that their child is describing, but the same description uttered by someone else's child may not be processed as quickly. While no studies directly test the effect of prior interaction on perspective-taking in children, our model may help explain the seemingly contradictory results in the field, particularly in the developmental literature. Our account predicts that children will struggle the most with unfamiliar partners, but that accuracy will improve over multiple interactions, and these predictions are supported by some studies (e.g., Krauss & Glucksberg, 1977). On the other hand, our account predicts that children should succeed in communicating with familiar partners. In line with this prediction, Leung et al. (2020) find that even 4-year-olds can succeed at a matcher-director task involving novel shapes when interacting with their parents. Further studies could contrast children's interactions with familiar and unfamiliar adults, as well as the effects of continued interaction (i.e., child interacts with the same adult throughout the session) on communicative perspective-taking.

Concluding Remarks and Suggested Future Directions

In the final portion of this chapter (and dissertation), I consider the findings from Chapters 1-3 from the perspective of the proposed online computation account and provide some future directions for testing the proposed account. In Chapter 1, I found that parents maintain a highly-specific model of their children's vocabulary knowledge, and spontaneously recruit that model to design speech for effective communication. In Chapter 2, I show that parents and children adapt to one another in various ways when coordinating with each other in a referential task, and show that production (but not comprehension) constraints may be an underlying cause of children's poor referential skill. Finally, in Chapter 3, I show that young children can make fine-grained judgments about a younger child's vocabulary knowledge, but they are unable to leverage those representations during real-time communication.

One of the key findings in Chapter 1 is that parents not only utilize their complex representations of their children's vocabulary during communication, they can also make real-time adjustments. This suggests that our interactions with others likely make use of both precomputed models of the other person, as well as online computations of the other person's perspective. Chapter 1 provides some insight into how these different types of computations interact. Beyond producing language tailored to their children based on prior knowledge, parents also adjusted their speech in real time. Specifically, if a child responded incorrectly on a trial where parents originally believed their child knew the animal, parents used longer referring expressions to describe that animal when it appeared again in the study. That is, parents did not merely rely on their extensive prior knowledge of their children, but made in-the-moment adaptations where necessary. While precise response times were not measured in this study, the online computation theory would predict slower response times for trials where parents made in-the-moment adjustments. Future studies could use speech onset times (or other measures of response time) to investigate how precomputation and online computation differentially effect language processing in communication.

In Chapter 2, I examined the conversational patterns of parents and children during a cooperative reference game. Similar to Chapter 1, I find evidence that parents leverage their knowledge of their children's developmental level to scaffold referential success. Parents of 4-year-olds spoke significantly more than parents of 6- and 8-year-olds, suggesting that parents are modulating the amount of information they provide, presumably in service of communicative success. Importantly, Chapter 2 also allows us to consider the models that children may have of their parents. Because children are interacting with a familiar partner (a caregiver), they already have expectations for the conversational dynamic. Further, based on their extensive prior interaction, children may be able to assume that their parents understand them. Indeed, both parents and children were very accurate at selecting the target during the study, indicating that both parents and children understood each other. Interestingly, I find that parents are more likely to provide referential expressions that ultimately become reused—that is, parents are the initiators of referential pacts. This suggests that children willingly take up the expressions used by their parents, rather than stubbornly continue with their own descriptions. It is possible that children, when faced with a difficult referential task, reduce their computational load by adapting to their parents. Because children are clearly very familiar with their caregivers, a remaining question is whether children are equally likely to adapt to an unfamiliar adult. Is children's willingness to reuse their parents' expressions a product of their many prior repeated interactions, or is the adaptation purely to minimize computational load?

In Chapter 3, I remove the supportive parental context to ask about children's models of others' lexical knowledge, and whether they can use those models during real-time communication. I find that even 4-year-old children have surprisingly accurate estimates of a younger child's vocabulary knowledge, and the ability to infer other's vocabulary develops significantly from ages 4 to 8. However, analysis of a partial sample of data indicates that children do not modify speech based on their beliefs about a younger child's vocabulary. Even more surprisingly, children do not adapt even when explicitly told about the younger child's knowledge/ignorance. What does this mean from the perspective of the online computation account? The account predicts that children would be more likely to make adaptations when knowledge information is given, because no real-time inferences need to be made. However, it is possible that the cognitive load required to generate an alternative description is still too high. An interesting future direction would be test the impact of live feedback and repeated interactions. That is, if young children receive real-time feedback that their conversational partner does not know something, will they make adjustments when they refer to the same object again? The online computation account predicts that with more repeated interaction and feedback, children should be more likely to adapt to their partners—this is because information can become pre-computed over time. Further, feedback can serve as an additional source of information, strengthening children's beliefs about their conversational partner's knowledge. Another possible direction would be to provide children with partners on more "extreme" ends of knowledge or ignorance, to test whether children are capable of any linguistic adaptation. For example, if children are introduced to an alien, who has never encountered any objects on earth before, how will children decide to describe various items? The online computation account predicts that children should be more likely to adapt in these "extreme" situations, because the relative computations are simple—the alien should not know any labels, as opposed to a child who might know some, but not other, labels.

The proposed online computation account should be tested empirically. An important area of study may be to conduct studies with similar paradigms for both adults and children to compare the two populations. Currently, developmental work on communicative perspective-taking often lacks theoretical grounding. Most studies are interested in *when* children can (or cannot) communicate effectively with others, rather than *why* they may succeed or struggle to take others' perspectives. By applying the proposed theoretical framework onto developmental studies, we can deepen our understanding of the developmental trajectory of communicative ability, as well as examine the factors underlying successful communication in both adults and children.

Future studies may wish to quantify the predictions made by the online computation account. The online computation account predicts that more repeated interactions should lead to greater facilitation of perspective-taking due to more information being pre-computed. However, empirical data will help clarify how each additional repeated interaction builds upon prior interactions, and the time-course at which in-the-moment computations become precomputations. One question we may ask is: how *granular* is the proposed online computation account? Can we find differences in communicative perspective-taking when comparing individuals who have had 10 repeated interactions as opposed to 5? A second interesting question is: when does information move from in-the-moment computation to precomputation? If information is presented at the beginning of an interaction, can it become pre-computed later within the same interaction? Indeed, a finding that multiple studies converge upon is that conversational partners become “better” at communicating over time (e.g., Hawkins, et al., 2021). For example, in referential settings, selection accuracy usually increases over time. This is in line with an online computation account, and further experiments designed specifically to compare perspective-taking (and not task fluency, etc.) over time can shed light onto how information becomes pre-computed and thus more easily retrieved for perspective-taking purposes.

Communication, more often than not, is successful—that is, people seemingly take one another’s perspectives with relative ease. A rich body of work has documented this phenomenon of perspective-taking in communication, and several theories have been proposed to capture this behavior. Existing frameworks focus on explaining empirical data in adults, and are often constrained to particular experimental paradigms. The online computation account proposed in this paper provides a comprehensive framework for understanding communicative perspective-taking in both adults and children. The key claim of this account is that perspective-taking is contingent upon *in-the-moment* (or online) computational load. Information that must be accounted for during online communication, such as the complexity of the visual environment or new information about a conversational partner, directly influences an individual's ability and propensity to take their partner's perspective. In contrast to in-the-moment computations, some information may be *pre-computed*. These precomputations can occur when conversational partners engage in repeated interactions, thus building pre-existing knowledge about one another. The differentiation between in-the-moment computations and precomputations is crucial because it has the potential to explain the incredible variability in empirical data. Rather than attributing all differences to paradigmatic differences or complexity of stimuli, the proposed account offers a novel and interesting explanation for why people succeed—and fail—to communicate in a variety of situations. Further, this account can be applied to both adults and children, and provides a framework both for evaluating current work and designing future studies.

Children begin to participate in conversations within the first few years of life, yet their communicative skills develop significantly over childhood. In this dissertation, I have shown that parents play an important role in scaffolding children's communicative development by adapting their language (at the word-level and discourse-level) to children's developmental level (Chapters 1 and 2). Further, while children may face challenges in generating informative descriptions of objects, they have sensible intuitions about even the specific words other people may know (Chapters 2 and 3). By focusing on the interactive contexts that children engage in and by investigating both the successes and struggles they face, this dissertation paints a holistic picture of children's communicative development. As children's own developing linguistic and cognitive skills support their emerging ability to engage in communicative adaptation, they also actively participate in rich interactive environments where input is adapted to them.

References

Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development*, *67*(2), 635–645.

Anderson, A. H., Clark, A., & Mullin, J. (1994). Interactive communication between children: Learning how to make language work in dialogue. *Journal of Child Language*, *21*(2), 439–463.

Asher, S. R., & Oden, S. L. (1976). Children's failure to communicate: An assessment of comparison and egocentrism explanations. *Developmental Psychology*, *12*(2), 132.

Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, 117–134.

Baer, C., & Friedman, O. (2018). Fitting the message to the listener: Children selectively mention general and specific facts. *Child Development*, *89*(2), 461–475.

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

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Beal, C. R. (1987). Repairing the message: Children's monitoring and revision skills. *Child Development*, 401–408.

Beal, C. R. (1988). Children's knowledge about representations of intended meaning. *Developing Theories of Mind*, 315–325.

Beal, C. R., & Belgrad, S. L. (1990). The development of message evaluation skills in young children. *Child Development*, *61*(3), 705–712.

Beal, C. R., & Flavell, J. H. (1982). Effect of increasing the salience of message ambiguities on kindergartners' evaluations of communicative success and message adequacy. *Developmental Psychology*, *18*(1), 43.

Beal, C. R., & Flavell, J. H. (1983). Young speakers' evaluations of their listener's comprehension in a referential communication task. *Child Development*, 148–153.

Bernstein Ratner, N. (1984). Patterns of vowel modification in mother–child speech. *Journal of Child Language*.

Birch, S. A., & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental-state attribution. *Psychological Science, 14*(3), 283–286.

Bjork, R. A., & Kroll, J. F. (2015). Desirable difficulties in vocabulary learning. *The American Journal of Psychology*, *128*(2), 241.

Blewitt, P. (1983). Dog versus collie: Vocabulary in speech to young children. *Developmental Psychology*, *19*(4), 602.

Bonawitz, E., Denison, S., Griffiths, T. L., & Gopnik, A. (2014). Probabilistic models, learning algorithms, and response variability: Sampling in cognitive development. *Trends in Cognitive Sciences*, *18*(10), 497–500.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning Memory and Cognition*, *22*(6), 1482–1493.

Branigan, H. P., Bell, J., & McLean, J. F. (2016). Do you know what I know? The impact of participant role in children's referential communication. *Frontiers in psychology*, *7*, 213.

Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialog. *Topics in Cognitive Science*, *1*(2), 274-291.

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*(2), B33–B44.

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The bank of standardized stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PloS One*, *5*(5), e10773.

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of standardized stimuli (BOSS) phase II: 930 new normative photos. *PloS One*, *9*(9), e106953.

Brown, R. (1958). How shall a thing be called? *Psychological Review*, *65*(1), 14.

Brown-Schmidt, S. (2009). Partner-specific interpretation of maintained referential precedents during interactive dialog. *Journal of Memory and Language*, *61*(2), 171–190.

Brown-Schmidt, S., & Hanna, J. E. (2011). Talking in another person’s shoes: Incremental perspective-taking in language processing. *Dialogue & Discourse*, *2*(1), 11–33.

Brown-Schmidt, S., Yoon, S. O., & Ryskin, R. A. (2015). People as contexts in conversation. In *Psychology of learning and motivation* (Vol. 62, pp. 59–99). Elsevier.

Bruner, J. (1983). *Child’s talk: Learning to use language* (pp. 111–114). Norton.

Bruner, J. (1985). Child’s talk: Learning to use language. *Child Language Teaching and Therapy*, *1*(1), 111–114.

Clark, E. V. (2009). *First language acquisition*. Cambridge University Press.

Clark, E. V. (2018). Conversation and language acquisition: A pragmatic approach. *Language Learning and Development*, *14*(3), 170–185.

Clark, H. H. (1996). *Using language*. Cambridge university press Cambridge.

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. In *Advances in psychology* (Vol. 9, pp. 287–299). Elsevier.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.

Cooper, R. P., & Aslin, R. N. (1990). Preference for Infant-Directed Speech in the First Month

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology*, *65*(2), 171–237.

Datavyu Team. (2014). Datavyu: A video coding tool. *Databrary Project, New York University*. http://datavyu.org

Degen, J., Hawkins, R. D., Graf, C., Kreiss, E., & Goodman, N. D. (2020). When redundancy is useful: A bayesian approach to "overinformative" referring expressions. *Psychological Review*, *127*(4), 591.

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.

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition*, *11*(2), 159–184.

Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science*, *33*(4), 547–582.

Endress, A. D., Nespor, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*, *13*(8), 348–353.

Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, *25*(5), 1–54.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & others. (2007). *MacArthur-bates communicative development inventories: User's guide and technical manual*. Baltimore, MD: Brookes.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, *20*(1), 104.

Fitneva, S. A. (2010). Children's representation of child and adult knowledge. *Journal of Cognition and Development*, *11*(4), 458–484.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, *44*(3), 677–694.

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, *62*(1), 35–51.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, *2*(4), 1360–1383.

Ghrear, S., Fung, K., Haddock, T., & Birch, S. A. (2021). Only familiar information is a "curse": Children's ability to predict what their peers know. *Child Development, 92*, 54-75.

Glucksberg, S., & Krauss, R. M. (1967). What do people say after they have learned how to talk? Studies of the development of referential communication. *Merrill-Palmer Quarterly of Behavior and Development*, *13*(4), 309–316.

Glucksberg, S., Krauss, R. M., & Weisberg, R. (1966). Referential communication in nursery school children: Method and some preliminary findings. *Journal of Experimental Child Psychology*, *3*(4), 333–342.

Goldberg, A. E., & Ferreira, F. (2022). Good-enough language production. *Trends in Cognitive Sciences*.

Grieser, D. A. L., & Kuhl, P. K. (1988). Maternal Speech to Infants in a Tonal Language: Support for Universal Prosodic Features in Motherese. *Developmental Psychology*.

Haber, J., Baumgärtner, T., Takmaz, E., Gelderloos, L., Bruni, E., & Fernández, R. (2019). The PhotoBook dataset: Building common ground through visually-grounded dialogue. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 1895–1910.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, *49*(1), 43-61.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young american children.* Paul H Brookes Publishing.

Hawkins, R. D., Frank, M. C., & Goodman, N. D. (2020). Characterizing the dynamics of learning in repeated reference games. *Cognitive Science*.

Hawkins, R. D., Franke, M., Frank, M. C., Goldberg, A. E., Smith, K., Griffiths, T. L., & Goodman, N. D. (2022). From partners to populations: A hierarchical bayesian account of coordination and convention. *Psychological Review*.

Hawkins, R. D., Gweon, H., & Goodman, N. D. (2021). The division of labor in communication: Speakers help listeners account for asymmetries in visual perspective. *Cognitive Science, 45*(3), e12926.

Hockett, C. F., & Hockett, C. D. (1960). The origin of speech. *Scientific American*, *203*(3), 88–97.

Horton, W. S., & Gerrig, R. J. (2002). Speakers' experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language*, *47*(4), 589–606.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground?. *Cognition*, *59*(1), 91–117.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*(2), 151–195.

Hupet, M., Seron, X., & Chantraine, Y. (1991). The effects of the codability and discriminability of the referents on the collaborative referring procedure. *British Journal of Psychology*, *82*(4), 449–462.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343–365.

Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics*, *4*(1), 71-84.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, *11*(1), 32-38.

Koranda, M. J., Zettersten, M., & MacDonald, M. C. (2022). Good-enough production: Selecting easier words instead of more accurate ones. *Psychological Science*, *33*(9), 1440–1451.

Krauss, R. M., & Glucksberg, S. (1969). The development of communication: Competence as a function of age. *Child Development*, 255–266.

Krauss, R. M., & Glucksberg, S. (1977). Social and nonsocial speech. *Scientific American*, *236*(2), 100–105.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831–843.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 english words. *Behavior Research Methods*, *44*(4), 978–990.

Lachman, R., Shaffer, J. P., & Hennrikus, D. (1974). Language and cognition: Effects of stimulus codability, name-word frequency, and age of acquisition on lexical reaction time. *Journal of Verbal Learning and Verbal Behavior*, *13*(6), 613–625.

Leung, A., Hawkins, R., & Yurovsky, D. (2020). Parents scaffold the formation of conversational pacts with their children. In S. Denison M. Mack Y. Xu & B. C. Armstrong (Eds.), *Proceedings of the 42nd annual conference of the Cognitive Science Society* (pp. 1022–1028). Austin, TX: Cognitive Science Society.

Leung, A., Tunkel, A., & Yurovsky, D. (2021). Parents fine-tune their speech to children's vocabulary knowledge. *Psychological Science*, *32*(7), 975–984.

Liu, H.-M., Tsao, F.-M., & Kuhl, P. K. (2009). Age-related changes in acoustic modifications of mandarin maternal speech to preverbal infants and five-year-old children: A longitudinal study. *Journal of Child Language*, *36*(4), 909.

Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development*, *82*(6), 1778–1787.

Ma, W., Golinkoff, R. M., Houston, D. M., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development*, *7*(3), 185–201.

Majid, A., & Burenhult, N. (2014). Odors are expressible in language, as long as you speak the right language. *Cognition*, *130*(2), 266–270.

Majid, A., Roberts, S. G., Cilissen, L., Emmorey, K., Nicodemus, B., O'grady, L., Woll, B., LeLan, B., De Sousa, H., Cansler, B. L., et al. (2018). Differential coding of perception in the world's languages. *Proceedings of the National Academy of Sciences*, *115*(45), 11369–11376.

ManyBabies Consortium (2020). Quantifying sources of variability in infancy research using the infant-directed-speech preference. *Advances in Methods and Practices in Psychological Science*, *3*(1), 24–52.

Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development*, 986–992.

Masur, E. F. (1997). Maternal labelling of novel and familiar objects: implications for children's development of lexical constraints. *Journal of Child Language*, *24*, 427–439.

Matthews, D., Lieven, E., & Tomasello, M. (2007). How toddlers and preschoolers learn to uniquely identify referents for others: A training study. *Child Development*, *78*(6), 1744–1759.

Mervis, C. B., & Mervis, C. A. (1982). Leopards are kitty-cats: Object labeling by mothers for their thirteen-month- olds. *Child Development*, 267–273.

Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects in the comprehension of referring expressions. *Journal of Memory and Language, 49*, 201–213.

Moerk, E. L. (1976). Processes of language teaching and training in the interactions of mother-child dyads. *Child Development*, 1064–1078.

Murthy, S. K., Griffiths, T. L., & Hawkins, R. D. (2022). Shades of confusion: Lexical uncertainty modulates ad hoc coordination in an interactive communication task. *Cognition*.

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of Perspective-Taking Constraints in Children's On-Line Reference Resolution. *Psychological Science*, *13*(4), 329–336.

Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. A. Ferguson (Ed.), *Talking to children language input and interaction* (pp. 109–149). Cambridge University Press.

Nikolaus, M., Prévot, L., & Fourtassi, A. (2022). Communicative feedback as a mechanism supporting the production of intelligible speech in early childhood. *Proceedings of the 44th Annual Conference of the Cognitive Science Society*, 771–778.

Phillips, J. R. (1973). Syntax and vocabulary of mothers' speech to young children: Age and sex comparisons. *Child Development*, 182–185.

Robinson, E. J., & Robinson, W. P. (1977). Children's explanations of communication failure and the inadequacy of the misunderstood message. *Developmental Psychology*, *13*(2), 156–161.

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, *29*(5), 700–710.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*, 217–236.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*(5), 1762–1774.

Roy, B., Frank, M., & Roy, D. (2009). Exploring word learning in a high-density longitudinal corpus. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 31).

Ryskin, R., Stevenson, S., & Heller, D. (2020). Probabilistic weighting of perspectives in dyadic communication. *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society*.

Sidera, F., Perpiñà, G., Serrano, J., & Rostan, C. (2018). Why Is Theory of Mind Important for Referential Communication?. *Current Psychology (New Brunswick, N.J.)*, *37*(1), 82–97.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis* (Vol. 1). Cambridge University Press.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, *21*(2), 211–232.

Shatz, M., & Gelman, R. (1973). The development of communication skills: Modifications in the speech of young children as a function of listener. *Monographs of the Society for Research in Child Development*, *38*(5, Serial No 152), 1–37.

Shneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science*, *15*(5), 659–673.

Simpson, E. H. (1949). Measurement of diversity. *Nature*, *163*(4148), 688–688.

Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word–referent learning. *Trends in Cognitive Sciences*, *18*(5), 251–258.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174.

Snow, C. E. (1972). Mothers’ Speech to Children Learning Language. *Child Development*, *43*(2), 549–565.

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, *27*(4), 501–532.

Speer, J. R. (1984). Two practical strategies young children use to interpret vague instructions. *Child Development*, 1811–1819.

Tamis-LeMonda, C. S., Kuchirko, Y., Luo, R., Escobar, K., & Bornstein, M. H. (2017). Power in methods: Language to infants in structured and naturalistic contexts. *Developmental Science*, *20*(6), e12456.

Taylor, M., Cartwright, B. S., & Bowden, T. (1991). Perspective taking and theory of mind: Do children predict interpretive diversity as a function of differences in observers’ knowledge? *Child Development*, *62*(6), 1334–1351.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*(6022), 1279–1285.

Vlach, H. A., & Sandhofer, C. M. (2014). Retrieval dynamics and retention in cross-situational statistical word learning. *Cognitive Science*, *38*(4), 757–774.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard university press.

Walley, A. C., & Metsala, J. L. (1992). Young children's age-of-acquisition estimates for spoken words. *Memory & Cognition*, *20*(2), 171–182.

Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*, *24*(11), 2143–2152.

Wingfield, A. (1968). Effects of frequency on identification and naming of objects. *The American Journal of Psychology*, *81*(2), 226–234.

Wojcik, E. H., Zettersten, M., & Benitez, V. L. (2022). The map trap: Why and how word learning research should move beyond mapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, e1596.

Yoon, S. O., Jin, K., Brown-Schmidt, S., & Fisher, C. L. (2021). What's new to you? Preschoolers' partner-specific online processing of disfluency. *Frontiers in Psychology*, *11*, 612601.

Yurovsky, D. (2018). A communicative approach to early word learning. *New Ideas in Psychology*, *50*, 73–79.

Yurovsky, D., Yu, C., & Smith, L. B. (2012). Statistical speech segmentation and word learning in parallel: Scaffolding from child-directed speech. *Frontiers in Psychology*, *3*.

Zettersten, M., & Lupyan, G. (2020). Finding categories through words: More nameable features improve category learning. *Cognition*, *196*, 104135.