

ABSTRACT

Title of Dissertation: NAVIGATING LINGUISTIC COMPLEXITY
IN ACQUISITION:
FROM SYNTACTIC DEPENDENCIES
TO REFERRING EXPRESSIONS

Jack Yuanfan Ying
Doctor of Philosophy, 2025

Dissertation Directed by: Professor Jeffrey Lidz
Department of Linguistics

Professor Alexander Williams
Department of Linguistics

This dissertation explores how young learners navigate complex linguistic structures and form generalizations under two challenging conditions: sparse input and unclear input. I focus on two case studies, the comprehension of instrumental *with*-questions and the production of definite *the*-expressions.

The first case study examines how 15-month-olds comprehend instrumental *wh*-questions with a stranded preposition *with* (*What did she break the stick with?*). I ask whether infants' resolution of these prepositional object gaps in *wh*-questions will

fail due to usage factors—infrequent question types and verb-based expectations—or if they will succeed via detecting a missing argument, signaled by the complement-selecting preposition *with*. Despite input sparsity and unfavorable verb expectations, infants responded correctly to these questions, suggesting that they are pragmatically prompt in responding to *wh*-questions before fully acquiring their full syntax.

The second case study concerns the acquisition of definite NPs. Naturalistic data suggests children’s referential choices aligned closely with their mothers’, with low miscommunication and high guessability. An interactive elicited production task shows that even 3- to 4-year-olds adapted their referring expressions appropriately, indicating adult-like knowledge of *the*. Their reported overuse of *the* in prior production studies likely reflects task artifacts. Longitudinal video corpora reveal that anaphoric cues (*a ball...the ball*) in input to 14-month-olds were available but inconsistent, whereas situational cues were widely available. Crucially, input with variable uses at 14 months did not mislead learners’ later use of *the* at 20 months, implying that they integrate both anaphoric and situational cues in recognizing the familiarity presupposition of *the*.

Altogether, this dissertation contributes to our understanding of how language learning unfolds as a dynamic, inferentially rich process. Young learners are remarkably ingenious in generating structure and meaning from limited input at the very beginning. Furthermore, this work highlights the role of methodology in revealing linguistic competence in early development.

NAVIGATING LINGUISTIC COMPLEXITY IN ACQUISITION:
FROM SYNTACTIC DEPENDENCIES TO REFERRING
EXPRESSIONS

by

Jack Yuanfan Ying

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2025

Advisory Committee:

Professor Jeffrey Lidz, Chair/Advisor
Professor Alexander Williams, Co-Chair/Advisor
Professor Valentine Hacquard
Professor Ellen Lau
Professor Yi Ting Huang

© Copyright by
Jack Yuanfan Ying
2025

Any minute now, my ship is coming in
I'll keep checking the horizon
And I'll check my machine, there's sure to be that call
It's gonna happen soon, soon, oh, so very soon
It's just that times are lean

And you say, be still my love
Open up your heart, let the light shine in
Don't you understand
Oh, I already have a plan
I'm waiting for my real life to begin

On a clear day, I can see, see for a long way
On a clear day, I can see
See for a long way

—Colin Hay, *Waiting for My Real Life to Begin*

Acknowledgments

Being able to study linguistics at Maryland has truly been the greatest gift I've ever received. I still remember running all over my house, cheering like I had no upstairs neighbors, the day I received my admission letter. It is hard to believe how quickly time has flown—now I'm at the final stretch, about to complete this dissertation.

I'm deeply grateful to everyone I've met, whether over zoom or in person, who welcomed me into this vibrant community and made these five years, far from home, such a memorable and transformative experience.

First, to my beloved advisors, Jeff Lidz and Alexander Williams. I'm forever thankful for the advice and support they have offered along the way. They have both made me feel at home during all those advisor meetings, and I have to confess how I enjoyed the meetings, particularly when both of them were there. I cannot keep track of how many things I've learned from our conversations, but I remember always being so inspired by the visions they articulated when I was navigating the blooming buzzing confusion of doing language science. They have shown me what it's like to be a conversant and eloquent language scientist, and I feel so so lucky to have had that experience learning from them. These lessons will stay with me for miles and miles, as a baby researcher and beyond.

I also feel so fortunate to have worked with Valentine Hacquard. The course she and Jeff co-taught on the acquisition of semantics sparked my interest in *the*-puzzle, a definite gamechanger for my research. Valentine's patience, insight, and kindness have meant so much to me, and I deeply enjoyed our chilling yet productive meetings. She has shown me what it's like to be a graceful and charismatic language scientist.

Ellen Lau is the person who kindled in me love for neuroscience. I remember signing up for her class on the Cognitive Neuroscience of Language and being so affected by her way of thinking throughout. The handwritten books she printed, the thought clouds she drew on the board, and the moments when she paused. I have taken neurolinguistic courses before, but her course got me on board for the first time. Her encouragement made me believe in myself and in venturing into neuroscience as an acquisitionist. She has shown me what it's like to be a sincere and imaginative language scientist.

Thank you to Yi Ting Huang for being on my committee as the Dean's Representative, for always bringing joy to the language science community and making it a real *Disneyland of language science*! I have always enjoyed the great vibes she had whenever bumping into her at the Language Science Center. Gosh, I would really miss this lively community.

I also want to thank Colin Phillips. His passion for psycholinguistics always got me so intrigued and so motivated to start a project, even at late night, when I could only see him over zoom during my first year of classes. I will never forget him being so caring when I first arrived at College Park, and so thoughtful for putting me in contact with Masato Nakamura when I had nowhere to stay due to house issues. He has shown me what it's like to be a devoted and caring language scientist.

To the incredible faculty at Maryland—Naomi Feldman, Philip Resnik, Howard Lasnik, Kate Mooney, Omar Agha, Aron Hirsch, Masha Polinsky, Bill Idsardi, Andrea Zukowski, Juan Uriagereka, Norbert Hornstein, Michael Isreal, John Scott, Bob Slevc, Alex Chabot, Shevaun Lewis, and many more—thank you for helping me navigate the unknown and for encouraging me, even when my ideas were still half-baked. I have also learned so much from Tonia Bleam and Peggy Antonisse, who made me see the beauty of teaching and inspired me to keep learning and improving myself to be an all-round, amiable language scientist. And special thanks to Kim Kwok, for always being there, always ready to help, and for bearing with me every time I stopped by her office with one more thing to figure out. She truly made our lives easier. Salute to the continuous support she provided to save me and everyone else from many logistical disasters—thank you, Kim!

My time at Maryland wouldn't have been complete without the incredible students here who made this place feel like home. Thank you to Tyler Knowlton, Mina Hirzel, Yu'an Yang, Hisao Kurokami, Adam Liter, Daniel Goodhue, Craig Thornton, Masato Nakamura, Alex Krauska, Jéssica Mendes, Nika Jurov, Maša Bešlin, Polina Pleshak, Ciaran Stone, Imane Bou-Saboun, Luisa Seguin, Xinchu Yu, Rosa Eun-Kyoung Lee, Clara Cuonzo, Leslie Famularo, Fedya Golosov, Zulfiyya Aghakishiyeva, Jingyi Chen, Joselyn Rodriguez, London Dixon, Sathvik Nair, Utku Turk, Sebastián Mancha, Mal Shah, Sarah Boukendour, Cassandra Caragine, Xiaoyu Yang, Matthew Shi, Carmen Tang, Tessa Culleton, Jackson Confer, Imani Finkley, Sadira Lewis, Audrey Laun, and so many more. Thank you for the friendships, the hallway chats, and the shared journeys.

My work would not be possible without feedback from the Acquisition Lab. Thank you to Tara Mease, who was always there to listen and help—whether it was with lab matters or just navigating day-to-day challenges. I'm also thankful for my undergrad mentee Eli Herbst, who not only assisted on my baby *wh*-project and helped shoot stimuli films with Luisa Seguin, but later took my seminar and raised some truly amazing questions. Thanks also to past labbers Laurel Perkins and Anouk Dieuleveut for their pioneering work and encouragement, to my grad labmates Katherine Howitt, Lydia Quevedo, Allison Dods, Elizabeth Swanson, Will Zumchak, as well as undergrad and postbac labbers Danya Levy, James Burns, Ziqing Ji, Sterling Mullenix, Premvanti Patel, Chloë Miller, Saige Oechsli, Justine Yu, Avni Gulrajani, Sofia Bendaña, Robert Chappetta, Olivia La Fiandra, Nathalie Fernandez, Nathan Frenkel, Ada Bouthet, Alex Emmert, Garva Khandelwal, and many others—for your feedback and camaraderie.

My journey in Linguistics wouldn't have started without Xiaolu Yang, my advisor during my masters' program, who saw potential in me as a "raw" acquisition enthusiast and welcomed me into her lab. I'm also thankful for my collaborator at l'Université du Québec à Montréal, Rushen Shi, who offered invaluable guidance in experimental design and stimuli development for infant studies.

Of course, none of this would have been possible without support from my parents, Sufang Zhang and Haibing Ying. I remember defending my decision to pursue a PhD in Linguistics when they first heard it, and I'm grateful for their respect and support once they heard my conviction. They taught me so many things before I left home for the US—my mom gave me a crash course on making steamed buns, and my dad taught me how to drive and navigate bad traffic. But above all, they made me see that sometimes, you have to let go of what's close to go far.

I still can't quite believe that my five years at Maryland are coming to an end. It feels like waking up from a beautiful, lingering dream, as there's so many things here that are too good to be true, especially for me as a small town boy, traveling all the way to the other side of the globe for my first education away from home. Here, I've breathed in language science, molecules that slowly shaped me into the kind of inclusive, curious, and driven language scientist I aspire to be. This has become my home. And like all travelers, I'll have to move on to the next stop from time to time, and I'll return when life brings me back home.

Table of Contents

Dedication	ii
Acknowledgments	iii
Table of Contents	vi
List of Tables	ix
List of Figures	x
1 Introduction	1
1.1 Complex structures: Generalization beyond input	1
1.2 Comprehending sparse <i>wh</i> -questions in infancy	4
1.2.1 Identifying non-adjacent dependencies	5
1.2.2 How does the gap-filling strategy work?	9
1.3 Producing subtle referring expressions in early childhood	13
1.3.1 Use of <i>the</i> -expressions in natural production	14
1.3.2 Overcoming performance issues in elicited production	15
1.3.3 Learnability of <i>the</i> -expressions in early input	17
1.4 Overview	18
2 Transitivity guides question comprehension in early toddlers	20
2.1 Background	20
2.1.1 Prerequisites for identifying structural dependencies	22
2.1.2 Prerequisites for linking thematic roles	29
2.2 Study 1	32
2.2.1 Method	32
2.2.2 Results	33
2.3 Study 2	40
2.3.1 Method	40
2.3.2 Results	46
2.4 General discussion	51
3 Preschoolers' <i>the</i> -expressions in natural production	59

3.1	Background	59
3.1.1	The meaning of <i>the</i> -definites	61
3.1.2	The overuse literature	63
3.1.3	Reconsidering <i>the</i> -overuse	66
3.1.4	Hypotheses about children’s knowledge	68
3.2	Study 3	70
3.2.1	Method	70
3.2.2	Results	71
3.2.3	Discussion	76
3.3	Study 4	77
3.3.1	Method	77
3.3.2	Results	78
3.3.3	Discussion	80
3.4	Study 5	80
3.4.1	Method	81
3.4.2	Results	82
3.5	General discussion	86
4	Overcoming performance issues in elicited production	88
4.1	Background	88
4.1.1	Development of perspective taking	89
4.1.2	Other demands on encoding reference	92
4.2	Study 6	94
4.2.1	Method	94
4.2.2	Results	98
4.3	General discussion	103
5	How experience matters in learning <i>the</i> -expressions	108
5.1	Background	108
5.1.1	Sensitivity to discourse vs. contextual knowledge	114
5.1.2	Recognizing presuppositions	118
5.2	Study 7	121
5.2.1	Method	122
5.2.2	Results	126
5.2.3	Discussion	135
5.3	Study 8	139
5.3.1	Methods	139
5.3.2	Results	140
5.4	General discussion	150
6	Summary and outlook	154
6.1	Summary	155
6.2	Implications	157

6.2.1	Navigating syntactic dependencies	157
6.2.2	Navigating referring expressions	161
6.3	Broader implications	164
6.3.1	Early integration: Linking language to concept	164
6.3.2	The path to generalization: From input to target	167
6.3.3	Approaches: Probing competence through performance	170
6.4	Future directions	172
6.4.1	Developmental trajectory of question processing	172
6.4.2	Probing knowledge of discourse referents during childhood	175
6.4.3	Acquiring referring expressions in determiner-less languages	176
6.5	Conclusion	178
A	Data from Study 1	179
B	Stimuli for Study 2	183
C	<i>The</i> -misuse for 2-year-olds' corpora in Study 4	189
D	Stimuli for Study 5	194
D.1	Child determiner uses in 2-year-old corpora	195
D.2	Child determiner uses in 3-year-old corpora	214
D.3	Child determiner uses in 4-year-old corpora	235
D.4	Mother determiner uses in 2-year-old corpora	257
D.5	Mother determiner uses in 3-year-old corpora	277
D.6	Mother determiner uses in 4-year-old corpora	299
E	Stimuli for Study 6	324
E.1	Meeting Kermit	324
E.2	Introducing the setup	325
E.3	Practice trials	325
E.4	Test trials	328
E.5	Thanks and goodbye	332
	Bibliography	333

List of Tables

2.1	Verb-preposition rate in declaratives vs. <i>what</i> -questions	35
2.2	Proportion of instrumental <i>with</i> -specification by verb	39
2.3	Structure of Experimental Trial, Study 1	42
2.4	Test verbs: Counts and percentage uses with direct objects in the Brent corpus on CHILDES	45
3.1	Misuse of definite NPs for referent introduction in two- to five-year-olds	63
3.2	Revised rate of <i>the</i> -definites (excluding fragment NPs)	75
5.1	Inferring the presupposition of <i>the</i> -expressions through different uses	109
5.2	Basic information of children selected from the NewEngland corpus in Study 7	122
5.3	Basic information of children selected from the NewEngland corpus in Study 8	140
A.1	Proportion of <i>what</i> -questions by Education Level	179
A.2	Distribution of DO vs. PO <i>what</i> -questions (with subject questions removed)	181
A.3	Transitive vs. intransitive uses of preposition <i>with</i> across corpora . .	182
A.4	One sample t-tests: Proportion of target looking above chance level .	182
E.4	List of referential contexts in Study 6	328

List of Figures

1.1	Input for <i>with</i> -questions	2
1.2	Input for <i>the</i> -expressions	2
2.1	Proportion of Question Type by Corpus	34
2.2	Average proportion of gap by preposition	35
2.3	Heatmap of preposition frequency by Corpus (occurrences marked within cells)	37
2.4	Proportion of gap by preposition across corpora (proportion of gap marked within cells)	38
2.5	Images of an experimental video for a BREAKING scene	43
2.6	Image of a test phase for the BREAKING scene	43
2.7	Infants' proportion of target looks by delimited Window and Condition (the first vertical line in each facet denotes the offset of the auditory stimuli)	47
2.8	Average proportion of target looks to the instrument in the Wh condition compared to the Control condition	48
2.9	Average proportion of target looks to the instrument across three delimited windows in the Wh vs. Control condition	49
2.10	200 millisecond epochs from 1500 msec to 2000 msec after the offset of the questions in different windows	50
3.1	Rates of misusing the in previous elicitation studies (blue dotted line for children's mean and orange for adults')	67
3.2	Overall proportion of singular <i>the</i> -definites by age	71
3.3	Overall proportion of singular <i>the</i> -definites by 2-year-old corpus	72
3.4	Proportion of singular <i>the</i> -definites across clause types (S for statements/declaratives vs. Q for questions/interrogatives) by age	73
3.5	Proportion of singular <i>the</i> -definites across syntactic positions	74
3.6	A sample test trial during Study 4	83
3.7	Rate of correct guesses for definites by speakers and child age	84
3.8	Rate of correct guesses for indefinites by speakers and child age	84
3.9	Adults' error rates for definites by items, speaker, and child age	85
4.1	Experimental setup for Study 6	95

4.2	Rate of definites for Question 1 (adults vs. children)	98
4.3	Rate of definites for Question 2 (adults vs. children)	99
4.4	Rate of definites for unique referents by age in days (dashed line separates the data points for 3-year-olds and 4-year-olds to indicate group distinction)	100
4.5	Rate of definites for non-unique referents by age in days (dashed line separates the data points for 3-year-olds and 4-year-olds to indicate group distinction)	100
4.6	Percentage of NP forms for Question 1 by Group and Trial Type . . .	101
4.7	Percentage of NP forms for Question 2 by Group and Trial Type . . .	102
4.8	Percentage of misuse of <i>the N</i> by individuals for non-unique referents	102
4.9	Percentage of <i>the N</i> errors for unfamiliar referents across studies . . .	104
4.10	Percentage of <i>the N</i> errors for non-unique referents across studies . .	104
5.1	Proportion of use in first vs. subsequent mentions for ‘the’	128
5.2	Proportion of ‘the’ vs. ‘a’ by Discourse Status of the referent	129
5.3	Proportion of ‘the’ vs. ‘a’ by Discourse Status of the referent across child-parent dyads	130
5.4	Proportion of ‘the’ vs. ‘a’ by Discourse Status of the referent across child-parent dyads	131
5.5	Proportion of ‘the’ vs. ‘a’ by Discourse Status of the referent (filtering predicative uses) across child-parent dyads	132
5.6	Proportion of ‘the’ vs. ‘a’ in chain1 vs. chain2	133
5.7	Proportion of ‘the’ in chain1 vs. chain2	134
5.8	Average proportion of ‘the’ usages in chain1	135
5.9	Proportion of ‘the’ usages in chain1 across dyads	136
5.10	Average proportion of different ‘the’-usages at 14 months vs. 20 months	141
5.11	Proportion of hearer-old uses of ‘the’ at 14 months vs. 20 months . .	142
5.12	Count of determiner types (<i>the</i> vs. <i>a</i>) produced by 20-month-olds in different corpora	143
5.13	Use of ‘the’ by child 6 at 20 months (time: 0’57”)	144
5.14	Use of ‘the’ by child 16 at 19 months (time: 2’31”)	145
5.15	Use of ‘a’ by child 7 at 19 months (time: 8’40”)	147
5.16	Use of ‘the’ by child 7 at 19 months (time: 12’33”)	148
A.1	Proportion of DO vs. PO <i>what</i> -questions by Corpus	180
B.1	Images of an experimental video for an OPENING scene	183
B.2	Image of a test phase for the OPENING scene	183
B.3	Images of an experimental video for a CLOSING scene	184
B.4	Image of a test phase for the CLOSING scene	184
B.5	Images of an experimental video for a BREAKING scene	185
B.6	Image of a test phase for the BREAKING scene	185

B.7	Images of an experimental video for a CLEANING scene	186
B.8	Image of a test phase for the CLEANING scene	186
B.9	Images of an experimental video for a FIXING scene	187
B.10	Image of a test phase for the FIXING scene	187
B.11	Images of an experimental video for a TURNING scene	188
B.12	Image of a test phase for the TURNING scene	188
D.1	Instructions to participants before the experiment	194

Chapter 1: Introduction

1.1 Complex structures: Generalization beyond input

Children’s direct experience is often too limited by itself to determine the target of linguistic knowledge out of numerous possibilities, yet they acquire their language with striking speed and precision. This paradox—known as the Poverty of the Stimulus problem (Chomsky, 1965)—has not only revolutionized cognitive science but also fueled generations of language scientists dedicated to understanding how, from the very beginning of life, children transform fragmented input into a rich, rule-governed linguistic system. What underlying mechanisms enable young learners to achieve this feat? On one front, they must contend with sparse data, recognizing structures that rarely surface yet remain part of the system. An example of this is *with*-questions, whose input is shown in Figure 1.1.¹ On another, they face unclear input, where some forms are abundant but conditioned by subtle rules hidden in a speaker’s mind. This is illustrated by the case of *the*-expressions, as shown in Figure 1.2. Navigating these complexities—extracting stable patterns amid both sparsity and uncertainty of input—is no trivial task, especially in early development, when learners’ linguistic knowledge and cognitive resources are still developing.

¹Here, each black circle represents the total amount of utterances a baby hears in a month, and each dot represents a data point relevant for the target structure. This data is based on input to infants from 9 to 15 months (cf. Brent and Siskind, 2001). Figure 1.1 shows the input for questions ending with the preposition *with* (e.g., *What are we gonna play with?*). A learner hears them once every two days, assuming 5000 utterances a day (27/155432). Figure 1.2 shows the input for *the*-expressions (e.g., *the ball*). A learner hears them once every minute (12839/155432).

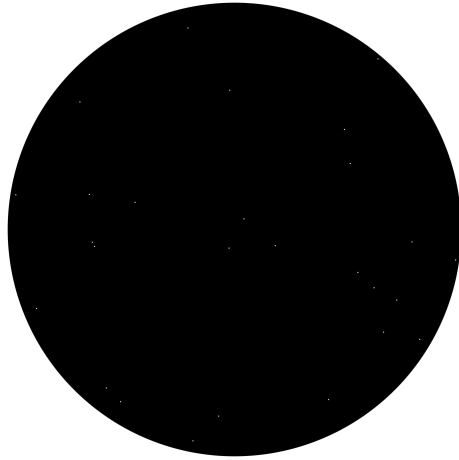


Figure 1.1: Input for *with*-questions

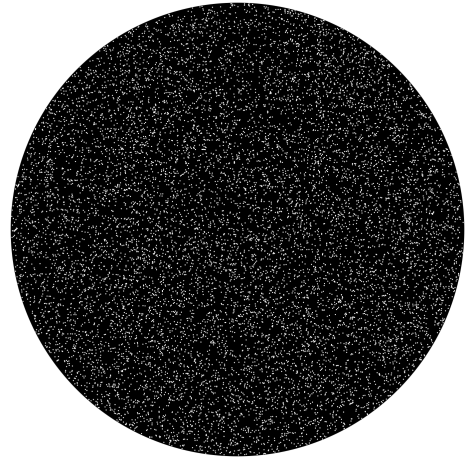


Figure 1.2: Input for *the*-expressions

This dissertation seeks to shed light on whether young children can acquire structures and meanings despite sparse and unclear input, and crucially, what underlying mechanisms allow them to bridge the gap between limited experience and the expansive knowledge defining human language.

For sparse input, I focus on how 15-month-old infants comprehend instrument questions with a stranded preposition *with* (e.g., *What did she break the stick with*). Prior studies suggest that infants are sensitive to the transitivity of verbs and use transitive verbs to uncover their following gap—a missing argument—in a direct object question. Do the selectional properties of linguistic categories in general, including prepositions, help guide young learners in question understanding? Are these learners' question understanding largely constrained by questions available in their input, such that they would become clueless when presented with these sparse prepositional object questions? Are their correct responses to questions contingent on their lexical expectations about verbs in particular?

Using a preferential-looking experiment, I find that 15-month-olds can resolve such *with*-questions. They succeed in interpreting these questions, even though these

instrument questions are sparse in their input, and despite the fact that the verbs chosen for the experiment typically do not occur with *with*-phrases, even in declarative clauses. Given prior research (Perkins and Lidz, 2021) showing that 15-month-olds do not yet represent *wh*-dependencies, our findings suggest that infants hold onto the partial structural cue, a missing argument after preposition *with*, in resolving the question.

For unclear input, I explore how 3- to 4-year-olds produce *the*-expressions. While these definite descriptions are abundant in input, their meanings can be obscure for learners, given the hidden presuppositions underlying variable referential contexts. Preschoolers have been reported to overuse *the*-expressions in contexts where the contextual requirements of *the* are not satisfied (Maratsos, 1976; Schafer and de Villiers, 2000; van Hout et al., 2010). Did these learners misuse *the*-expressions in prior studies due to deficient knowledge, or did the experimental setups inflate their errors? Do they consider conversational dynamics in choosing the right referential form in real time with a better controlled task design?

I assess young children’s production of these definite descriptions in natural and elicited settings. Contrary to previous claims that children overuse *the*-expressions, I show through a series of studies that children are highly adult-like in their use of singular *the*-expressions. In natural production, children do not overuse *the* compared to their caregivers across contexts, with minimal errors; their determiner choices (*the* vs. *a*) are as predictable as those of adults’. In elicited production, 3- to 4-year-olds adapt referring expressions to referential contexts as adults do, when extraneous task demands are minimized.

To explore the learnability of *the*-uses across toddlers, I examine a longitudinal video corpus of different caregiver-toddler pairs. *The*-definites require unique identifiability of intended referents, which can be established through discourse familiarity

(i.e., anaphorically) or contextual familiarity. I show that learners relying solely on anaphoric cues with noun repetition (e.g., *a ball...the ball*) may not reliably identify the familiarity presupposition of *the*-definites. Rather, they need to combine anaphoric and situational cues to arrive at an adult-like understanding of definite descriptions.

Finally, the dissertation discusses how findings from these two lines of inquiry contribute to our knowledge of possible learning mechanisms that bridge the gaps between input and linguistic knowledge.

1.2 Comprehending sparse *wh*-questions in infancy

Sparse structures in the input present the first challenge this dissertation will examine with young learners. This is especially the case with *wh*-questions, as they can take various forms, many of which are unheard of within a young learner’s limited experience.

Comprehending *wh*-questions in a language like English is non-trivial for infants. Structurally, they involve non-local movement-type dependencies between the *wh*-phrase/filler and the gap. For instance, for an object *wh*-question such as *What did she break _____*, young learners will have to recognize that *what* functions as the syntactic argument of *break*, despite being in a non-canonical position. This means that young learners have to start somewhere in finding *wh*-dependencies. They can, for instance, start with the gap by noticing how the direct object of the verb is displaced from its canonical position.

In recent studies (Gagliardi et al., 2016; Perkins and Lidz, 2020), it has been suggested that the argument structure of verbs can assist 15-month-old infants in comprehending object *wh*-questions and relative clauses without recognizing the non-

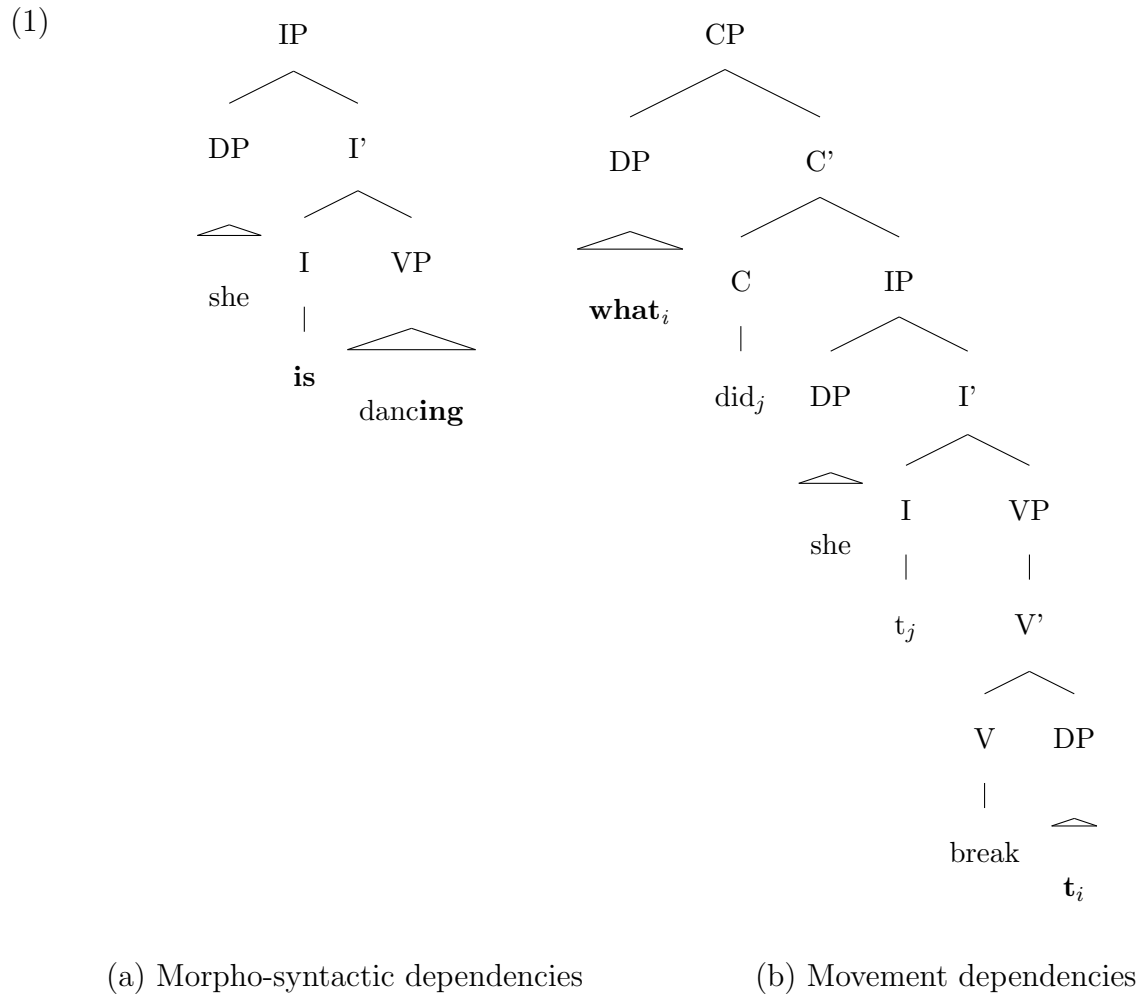
local dependencies involved. Infants may have succeeded in these tasks by noticing the gap following the verb and linking it to an unmentioned, corresponding event participant.

What remains unclear is how broadly infants' *gap-filling strategy* applies across different types of questions. Is the strategy guided by the missing argument signaled by the transitivity of its preceding category—a selectional requirement for an immediate complement—meaning that a gap following the preposition *with* should be just as effective for the strategy to work? Do questions with prepositional object gaps present a challenge for comprehension, given the sparsity of the question type? Since verbs serve as the head of a predicate and introduce core arguments, they may also play a key role in successful gap-filling. Are successful responses to the gap site after the preposition *with* contingent on learners' expectations about verb subcategorization frames? To explore these questions, we must first review the types of non-local dependencies learners represent in early development and consider potential learning mechanisms underlying their recognition of these syntactic dependencies.

1.2.1 Identifying non-adjacent dependencies

All natural languages feature grammatical rules that involve dependencies between non-neighboring elements. In English, for instance, this non-adjacent dependency (or discontinuous dependency) can take various forms. The first type can be found in *morpho-syntactic dependencies*, holding between two functional elements, as with an auxiliary *is* and verb inflection *-ing* (e.g., ***is dancing***). Although this dependency is morphologically non-adjacent, it is structurally local in nature, forming a sisterhood relationship within the syntactic tree, as illustrated in (1a). The second type can be found in *movement dependencies* or *filler-gap dependencies*, holding between a

displaced element or phrase and a variable it is coindexed with, as with a *wh*-phrase, i.e., the filler, and a gap, (e.g., **What** *did she break* ____). Different from morpho-syntactic dependencies, movement dependencies are structurally non-local, as shown in (1b).



Infants' ability to recognize morpho-syntactic dependencies is in place from early on. It was first observed in 4-month-old German-learning infants, who after familiarization with 2 morpho-syntactic dependencies (each with 32 variations for the

intervening element) in Italian, exhibited differentiated brain responses to grammatical vs. ungrammatical cases during test (Friederici et al., 2011). However, without familiarization, only English-learning 18-month-olds—not 15-month-olds—succeed in detecting morpho-syntactic dependencies, and only when there are no more than three intervening syllables (Santelmann and Jusczyk, 1998). This led to the proposal that infants’ ability to recognize such morpho-syntactic dependencies may be limited by their processing space. Later research (Hohle et al., 2006) demonstrated that infants can handle longer dependencies as long as they can assign syntactic structures to the intervening materials. Together, these findings suggest that enough instances with somehow-parsable intervening elements support infants to recognize morpho-syntactic dependencies.

In contrast, infants’ ability for recognizing movement dependencies is attested at a later stage. Existing studies (Gagliardi et al., 2016; Perkins and Lidz, 2020, 2021) suggest that argument structure is a precondition on identifying movement dependencies. While infants are sensitive to verb argument structures (Jin and Fisher, 2014) and respond correctly to object *wh*-questions at around 15 months (Gagliardi et al., 2016; Perkins and Lidz, 2020), they do not yet represent filler-gap dependencies in *wh*-questions until 18 months of age (Perkins and Lidz, 2021). This further suggests that argument structure may play a critical role in the acquisition of movement dependencies. Infants detect missing arguments at gap sites in the first place, before they can draw inferences about the discontinuous co-occurrence between the sentence-initial *wh*-phrase and the sentence-final gap it leaves behind.

Infants’ knowledge about verb transitivity, hence, becomes crucial to resolving the puzzle of learning movement dependencies. Perkins and Lidz (2020) found that only 15-month-olds with high vocabulary responded correctly to *wh*-questions and relative clauses, which suggests that knowledge of a verb’s argument structure may

facilitate infants’ question understanding. Indeed, there is converging evidence from children’s bootstrapping mechanisms for word learning (Gleitman, 1990), suggesting that toddlers around this age actively infer the meanings of novel verbs through observing their distribution in clauses (Fisher et al., 2010; Naigles, 1990) and linking thematic roles in their conceptual space to respective arguments in sentences (Perkins et al., 2024). This ability to infer verb meanings based on their clausal distribution (transitive vs. intransitive) has been observed in infants as young as 15 months (Jin and Fisher, 2014). This points to one attractive possibility: infants before 18 months of age may still be acquiring the argument structure of verbs, so as to gather evidence about the locally missing argument after transitive verbs in object *wh*-questions, before arriving at adult-like generalizations about movement dependencies.

Viewing from all these findings, learners may need to employ different strategies in uncovering different types of non-adjacent dependencies. Timeline-wise, more knowledge of argument structure seems to be needed for infants to detect the gap in argument questions before acquiring the movement dependencies involved. Since only high vocabulary guarantees successful performance at 15 months (Perkins and Lidz, 2020), it is possible that infants at this age might rely heavily on the argument structure of verbs in particular to correctly respond to missing arguments in *wh*-questions. This raises an important question: does the argument-missing category matter for infants’ *gap-filling strategy* to work? The answer to this question will enrich our understanding of what learning mechanism is essential for uncovering movement dependencies.

1.2.2 How does the gap-filling strategy work?

The studies discussed above leave one crucial question unanswered: Does infants' successful question comprehension at 15 months reflect a general ability to recover missing structures? Specifically, what mechanisms underlie their ability to interpret questions with missing arguments?

The transitivity-driven hypothesis suggests that infants' gap-filling strategy is guided by a broader, category-general structural cue: selectional requirements. Here, what matters is whether a word—be it a verb or a preposition—is transitive and thus selects for a complement. In this view, both transitive verbs and prepositions can trigger the expectation of an additional argument, so the detection of a locally missing argument after either category would guide infants' search for the unmentioned event participant that conceptually matches the missing argument at the gap site. Consequently, learners should be just as sensitive to a transitive preposition like *with* as they are to a transitive verb, when it comes to detecting a missing argument and recovering its thematic role.

However, infants' successful question comprehension in prior research (Gagliardi et al., 2016; Perkins and Lidz, 2020) may be contingent on usage factors. For instance, frequent questions with direct object gaps in the input may contribute to learners' successful comprehension. Learners may be better at resolving gaps that they frequently encounter in questions in the input. If a certain type of *wh*-questions—say, one involving a prepositional object gap—is rare, infants may struggle to resolve the gap during comprehension. This is very likely, as the prepositional object gaps they contain occur far less frequently in questions than direct object gaps do in both child- and adult-directed corpora (Atkinson et al., 2018).

Alternatively, infants may have succeeded in previous studies, given the special

role of verbs in particular in guiding their gap-filling strategy. As the head of the clausal predicate, transitive verbs play a central role in linking subjects/agents and objects/patients as their satellite arguments, while prepositional *with*-phrases are adjuncts to verb phrases, introducing optional instrumental arguments. Indeed, prior research suggests that learners recruit verb information early in a sentence and make verb-based predictions that override prepositions in the upcoming structure (Lidz et al., 2017). Infants are likely to treat prepositional phrases—such as *with*-phrases—as peripheral. As a result, a prepositional object gap in an instrumental *with*-question may be overlooked.

Instrumental *with*-questions thus provide a critical testing ground for the transitivity-driven hypothesis. Imagine a learner who has encountered sentences like (2) but has never been exposed to sentences like (3).

(2) She broke the stick.

(3) What did she break the stick with?

Gap-filling with (3) is more likely to fail, if learners’ success in comprehending questions at this age is contingent on usage factors. First, the sparsity of prepositional object gaps in learners’ input would make it difficult for infants to deploy the interpretive strategy. Second, the verb *break*’s core arguments (e.g., *she* and *the stick*) are already satisfied in the clause. This might downplay the salience of any additional missing argument, especially for one inside a *with*-phrase that is structurally optional for the subcategorization frame of verb *break*.

However, the transitivity-driven hypothesis predicts success in gap-filling. If infants are attuned to the transitivity of *with*, then its presence—being stranded—should signal a missing argument and cue them to search for the corresponding event

participant that fulfills its thematic role. That is, infants would treat the stranded preposition as a robust structural signal prompting the search for an instrument, even in the absence of prior exposure to this specific question form.

The transitivity-driven hypothesis is particularly desirable both for its developmental feasibility and generalization power. First, from a developmental perspective, infants are remarkably sensitive to functional elements and their syntactic dependencies early in development (for a review, see Shi, 2014). They can phonologically distinguish functional elements (e.g., *the*) from lexical elements (e.g., *mug*) from birth (Shi et al., 1999) and use functional elements to segment words in their first year (Shi and Lepage, 2008). Around 12 months, they start to categorize lexical elements by using nearby functional elements as syntactic anchors (Babineau and Christophe, 2022; Mintz, 2006). Crucially, Lidz et al. (2017) found that by 16 months, learners already interpret a novel noun phrase following *with* as the instrument of an action described in a sentence. This suggests that infants at this age can assign an instrument role to NPs following the preposition *with*.

Second, from a typological standpoint, a transitivity-driven strategy offers a powerful mechanism for generalization across structures. It could help infants not only interpret structures with knowledge of local dependencies but also infer what kinds of syntactic gaps their language permits—especially in movement dependencies. One key cross-linguistic distinction concerns preposition stranding. While Germanic languages like English allow prepositions to be stranded in *wh*-questions, the vast majority of the world’s languages do not (Merchant, 1999). For instance, while English allows the preposition to remain in situ in *wh*-questions (e.g., *what did she break the stick with?*), languages like French require pied-piping (Ross, 1967), in which the entire *with*-phrase must be fronted.

(4) French

- a. Avec quoi a-t-elle cassé le bâton? (With what did-she break the stick?)
- b. *Quoi a-t-elle cassé le bâton avec? (What did-she break the stick with?)

Furthermore, the implications of this cross-linguistic variation extend beyond *wh*-questions. Whether a language allows preposition stranding can correlate with allowed structures in other movement environments such as topicalization, relativization, and comparatives (Merchant, 1999).²

Taken together, a transitivity-driven approach is not only developmentally plausible but may also allow learners to use the distribution of arguments, whether headed by verbs or prepositions, to infer from early on whether their language allows prepositional gaps—knowledge that goes beyond specific lexical items and extends to broader grammatical patterns. For English-learning infants, this would support the early recognition of prepositional gaps as licit structures, enabling them to easily process sentences with stranded prepositions when they arise in the input, despite their low frequency in early input.

In Chapter 2, I investigate the underlying mechanism behind infants’ gap-filling strategy. I present preliminary evidence from 15-month-olds that supports the transitivity-driven hypothesis, shedding new light on how young learners navigate syntactic struc-

²It has been shown in Merchant (2000) that whether preposition stranding is allowed can also have consequences for sluicing. For instance, in Greek—a language that disallows preposition stranding in *wh*-movement ((i))—sluiced constructions also require the preposition to be overtly realized ((ii)).

(i) *pjon miluse i ana me
who was.talking the Anna with
‘Who was Anna talking with?’

(ii) i ana miluse me kapjon ala en ksero *(me) pjon
the Anna was.talking with someone but not know-1SG with who
‘Anna was talking to someone, but I don’t know who.’

tures in face of sparse input.

1.3 Producing subtle referring expressions in early childhood

Unclear input for a structure presents another challenge for young learners. For instance, appropriate use of the most frequent English word, the definite article or determiner *the*, is notoriously challenging for children (Maratsos, 1976; Wexler, 2011) and second-language learners (Ionin, 2003).

The meaning of *the*-phrases is subtle, as one needs to examine various referential contexts to determine its core meaning. This has led to a long history of debate on *the*-meaning. Consider usages of *the* in the following examples.

- (5) Do you see *the mug*? (with a single mug in the room)
- (6) Jason picked a mug from his mug collection. He then filled *the mug* with some freshly brewed coffee.

In these examples, an adult speaker can combine the determiner *the* with a noun *mug* to refer to either a unique mug in a situation (5) or a familiar mug in a discourse (6). That is, for an intended referent mug, speakers signal their presuppositions of its uniqueness (Russell, 1905) in (5) and its familiarity (Strawson, 1952; Heim, 1982) in (6) through the referential choice of a *the*-definite.

In acquisition studies, young children have been claimed to overuse *the*-expressions. Across numerous production tasks, they have been found to use definites when indefinites were expected for intended referents that were neither unique (Wexler, 2011) nor in the common ground (Maratsos, 1976; Schaeffer and Matthewson, 2005). This leads to the first possibility about young learners' knowledge of singular *the*-definites: they may possess deficient knowledge either with the linguistic requirements or with

speaker-listener reasonings. In this view, this deficiency in knowledge is the primary cause for children’s inappropriate uses.

What is largely glossed over in prior research is the question of whether young children’s production errors may arise from experimental artifacts. This is likely, as different methods lead to varying results (van Hout et al., 2010; De Cat, 2015). Children’s behavior was adult-like when they were tested for comprehension (Syrett et al., 2010; Choi et al., 2018). This points to the second possibility: young learners have an adult-like knowledge but have trouble deploying it in artificial production setups.

In summary, producing referring expressions can be subtly influenced by production setups, as children’s performance varies across studies. Is their knowledge deficient or adult-like? Is *the*-overuse real? To answer these questions, I will first present existing views on the observed production errors and raise the questions to be addressed in different chapters of the dissertation.

1.3.1 Use of *the*-expressions in natural production

A common, if not dominant view on children’s production errors with *the*-phrases is that their knowledge is deficient. Some attribute the root of this deficiency to their non-adult-like reasoning: Either children are egocentric, meaning that they are less likely than adults to correctly estimate what’s on the mind of their interlocutors (Maratsos, 1976; Karmiloff-Smith, 1979; Piaget, 1955), or they are immature in pragmatic reasoning skills (Schaeffer and Matthewson, 2005; van Hout et al., 2010), preventing them from using *the* in an adult-like way despite the correct meanings they assign to it. The difference is that the egocentric account highlights children’s internal deficit in reasoning about other minds in general, whereas the pragmatic account

focuses on children’s deficit in reasoning about specific situations. Others ascribe the deficiency to children’s non-adult-like linguistic knowledge, proposing that these errors may instead stem from children having the wrong meaning for *the*, perhaps failing to presuppose uniqueness (Wexler, 2011).

If the deficient-knowledge hypothesis is right, one clear prediction is that overuse would show up in children’s production, be it in natural or elicited settings. It would be easy to detect noticeable signs of deficiency, either due to children’s lack of perspective-taking skills or due to their wrong knowledge about the meaning of *the*.

In Chapter 3, I present three studies to examine whether young children do overuse *the*-phrases in natural production. For the first study on child corpus with 27 children, I look at the overall distribution of *the*-phrases in natural production on average and across individuals. For the second study on the same corpus, I measure the rate of *the*-driven miscommunications in natural production. For the third study with the Gleason’s corpus, I adapt the human simulation paradigm and ask adults to guess determiners used by children and their caregivers, given snippets of text-based conversations, so as to assess the adult-like-ness of children’s determiner choice. Through these studies, I provide evidence about whether children’s knowledge is deficient and leads to overuse of *the*-definites, even in natural production. Next, I discuss the kind of setup that we need to directly probe young children’s use of referring expressions in elicited production.

1.3.2 Overcoming performance issues in elicited production

While natural production provides valuable insights into children’s language use, it has certain limitations. First, as common ground is predictable in conversations between children and their caregivers, children do not have to take perspectives of the

listener on the spot in natural production. It might be easier for children to assess what their caregiver knows as opposed to what an experimenter knows, so natural production may not allow us to probe children’s capacity for assessing listener knowledge when producing referring expressions. Second, scenarios in natural production might be insufficient to expose underlyingly deficient knowledge. Children’s success with using singular *the*-phrases may be limited to mostly easy cases in natural production, so it is possible that in experimental settings targeting one particular part of their linguistic knowledge, they may fail to demonstrate sensitivity to the presuppositions of *the*-definites. Together, these limitations with natural production motivate us to further probe children’s knowledge of singular *the*-definites in elicited production.

If children’s knowledge is essentially adult-like, how can we go about probing this knowledge, without introducing potential experimental artifacts? Viewing from children’s varying performance across studies, two aspects of an elicited production experiment may boost production errors with *the*. The first is the inappropriateness of the elicitation context. One crucial difference between natural production and elicited production is the artificiality involved in elicited production. Unlike natural production, elicited production often lacks an interactive context, which could be less engaging for child participants to consider the perspectives of their listener or experimenter. The second is the unclear domain of reference. Keeping track of the type and the token of referents in the domain of reference can be challenging without appropriate visual support, which could lead to inaccurate assessment of referent uniqueness within the restricted domain.

In Chapter 4, I present a new experimental design that controls for the two possible sources of artifacts mentioned above. I refine the experimental setting to create a natural, interactive context and a clear domain of reference. The findings from this study will help us probe children’s knowledge of definite descriptions with a proper

setup and understand how children’s linguistic performance interacts with task setups.

1.3.3 Learnability of *the*-expressions in early input

Acquiring the correct understanding of when to use *the*-definites may not be a trivial task, especially for young learners. Recognizing the different presuppositions associated with *the* may require learners to combine different sources of cues. This is especially the case, as *the*-expressions are used for uniquely identifiable referents that are established either through prior mentions (i.e., linguistic/anaphoric cues; *a ball...the ball*), or through pertinent situational knowledge (i.e., non-linguistic/situational cues; contextual givenness).

Investigating the usage profiles of definite and indefinite descriptions in early input is crucial for us to hypothesize learning mechanisms at work in identifying presuppositions associated with singular *the*-expressions. Existing work (Sneed, 2007; Fisher and Tokura, 1995) focus on anaphoric cues for distinguishing definites and indefinites, namely that definites are more likely used for second-mentions than indefinites.

However, the usability of these anaphoric cues and their generalizability for inferring adult-like meanings is questionable. First, utterances go by quickly, and the portion of data infants can analyze, the intake (Lidz and Gagliardi, 2015; Omaki and Lidz, 2015), may not always include the beginning of a discourse. The second-mentions coded in a corpus may sometimes be registered as first-mentions by toddlers, given their fluctuating attentional state. Therefore, anaphoric cues, such as first- and second-mentions that annotators zero in on, may not be a veridical representation of the portion of data based on which young learners generalize about the meaning of *the*-definites. Second, anaphoric uses may not be dominant among other uses. This is a solid concern, as a great number of definite descriptions can be discourse-new,

thus non-anaphoric, even in adult corpora (Poesio and Vieira, 1997). In other words, general patterns observed from corpus data collapsed across multiple caregiver-child pairs may not reflect the varying input learners receive across scenarios. Little is understood about the variations in early input on definite descriptions and how that could influence the early output of learners.

In Chapter 5, I examine the early input that children receive by analyzing longitudinal video corpora from multiple caregiver-child pairs. Are surface anaphoric cues in distribution sufficient for young learners to preliminarily infer the meaning of *the*-definites in light of their developmental trajectory? Do learners need situational cues to recognize the familiarity presupposition at all? By focusing on the structure of the input over time, this chapter sheds light on how children may overcome the challenges of learning the meaning of *the* under variable input conditions. In doing so, it contributes to a broader understanding of how definiteness is acquired and what this reveals about the learnability of meanings in the face of unclear input.

1.4 Overview

In summary, this dissertation investigates early language development with sparse or unclear input and what learning mechanisms must be in place for learners to navigate different linguistic complexities given limited knowledge and cognitive resources.

This dissertation is structured as follows. In Chapter 2, I examine how 15-month-old infants tackle instrument questions with a stranded preposition *with*. I ask whether their successful resolution of argument questions reflects a general capacity for recovering missing arguments signaled by a preceding transitive category that selects for an immediate complement, or if it is contingent on usage factors. Our results show that infants can resolve these sparse questions with local structural cues

of transitivity, even though the missing argument follows the preposition *with*.

In Chapter 3 and 4, I investigate children’s production of *the*-expressions in natural and elicited settings. I show through different measures that preschoolers behave adult-like in producing *the*-expressions, either in daily conversations with caregivers or in an elicited production task with an improved setup. This suggests that children do not have linguistic or cognitive deficit in using proper referring expressions, and that their overuse of *the*-expressions claimed in the literature is most likely task-driven.

In Chapter 5, I ask the learnability of *the*-meaning across variable input conditions, using longitudinal videotaped corpora of caregiver-child interactions with children around 14 months and 20 months. While anaphoric cues—whether the referent has been previously mentioned—may not be consistently available in the usages of *the*, situational cues about whether the referent is contextually given are widely available to inform learners of the familiarity presupposition underlying *the*-expressions for first-mention referents. I further show that learners produced proper *the*-expressions at 20 months and understand when to use *a*-expressions instead. This suggests that learners do not rely solely on semantically less constraining anaphoric cues—they are sensitive to situations in which *the*-expressions are used.

These two case studies yield new insights into how young learners navigate complex structures with limited knowledge and cognitive resources. The findings reported in this dissertation highlight how learners can make linguistic generalizations about either syntax or semantics across linguistic and non-linguistic domains, despite unfavorable input conditions. They push us to reflect on choosing the relevant phenomena and the relevant methods to uncover the footprints of learning as human learners navigate different types of linguistic complexity.

Chapter 2: Transitivity guides question comprehension in early toddlers

2.1 Background

Learning *wh*-questions poses two interrelated challenges for young language learners: (1) identifying the underlying syntactic dependencies and (2) grasping the meaning upon hearing the structure. Take, for example, a *wh*-question like *What did she break?* Learners need to recognize the filler-gap dependency between the *wh*-phrase *what* (the filler) and the missing argument (the gap) following the verb *break*. In this case, *what* functions as a non-local argument of the verb *break*, meaning it is not adjacent to the verb in surface structure.

Representing this type of non-local syntactic dependencies can be particularly difficult for young learners. Research suggests that infants do not reliably show sensitivity to the complementary distribution of an object *wh*-phrase and a local object of a verb until around 18 months of age (Perkins and Lidz, 2020).

Despite challenges with fully parsing these questions, infants appear capable of understanding *wh*-questions before they fully grasp the syntactic dependencies involved. Studies have found that infants as young as 15 months can respond correctly to English *wh*-questions (Gagliardi et al., 2016; Perkins and Lidz, 2020). One proposed explanation is that infants rely on a *gap-filling* strategy. Rather than representing the full hierarchical structure of the sentence, they may detect a missing argument,

such as a missing direct object after a transitive verb, and infer the intended meaning for them to search for the event participant that corresponds to the missing argument. By this account, infants use local structural cues—particularly those informed by their knowledge of verb argument structure—to interpret *wh*-questions. Over time, this strategy may support the acquisition of the more complex, non-local dependencies involved (Gagliardi et al., 2016; Perkins and Lidz, 2020).

These findings suggest that by 15 months, learners have a basic understanding of what *wh*-questions are seeking, even if they have not yet mastered the syntactic dependencies involved. This early interpretive ability may explain, in part, why language learning proceeds so efficiently. By recognizing missing arguments—gaps after transitive categories like verbs—infants are able to roughly infer the intended meaning of a question, putting them in a favorable position to learn the associated grammatical structures.

However, an open question remains as to whether this gap-filling strategy represents a general ability for recovering missing arguments. Does the strategy extend to other types of argument questions, or is it specific to direct object *wh*-questions, where the gap follows a verb? In English, not all argument gaps occur in this position. For instance, *wh*-questions can also involve gaps that follow prepositions, such as in instrument questions like *What did she break the stick with?* Are infants at this age capable of recovering locally missing structures in these instrumental *with*-questions? This work seeks to fill this gap.

In this chapter, I first report a corpus study looking at the usage profiles of such prepositional object questions in input to infants at 9 to 15 months. The findings confirm the sparsity of such structures in the input, both in terms of their question type and the co-occurrence of early verbs with *with*-phrases. I then present an inter-modal preferential looking study designed to test 15-month-olds' comprehension of

instrumental *with*-questions, such as *What did she break the stick with?* In this type of question, the missing argument or gap follows the preposition *with* rather than a verb. The results show that 15-month-old infants—who are not yet sensitive to the full syntactic dependencies of *wh*-questions—are nonetheless able to comprehend these instrument questions, even though such forms are relatively rare in infant-directed speech.

Together, these findings suggest that 15-month-old infants can comprehend argument questions through detecting locally missing arguments, whether after verbs or prepositions. Importantly, their ability to resolve *wh*-questions is not contingent on usage factors like the frequency of the question type (Atkinson et al., 2018) or specific lexical expectations tied to particular verbs (Lidz et al., 2017). Instead, it appears that infants at 15 months already possess a more general capacity to use local structural information to infer the intended meaning of argument questions, supporting their broader syntactic development.

2.1.1 Prerequisites for identifying structural dependencies

Broadly, languages exhibit two types of syntactic dependencies: local dependencies, which occur between a head and its immediate complement, and non-local dependencies, which hold between two constituents that are structurally related but not adjacent in linear order. Local dependencies can sometimes appear as non-local elements on the surface due to intervening materials. For example, in English, there is a morpho-syntactic dependency between an auxiliary (e.g., *is*) and the form of the head of its VP complement (e.g., *-ing*), as in *is baking*. In contrast, non-local dependencies include those created by syntactic movement, such as filler-gap dependencies in *wh*-questions. In *What did she break?*, the *wh*-phrase *what* originates in the ob-

ject position of the verb *break* but is fronted, leaving behind a “gap”—an empty or unpronounced category required by the grammar.

To acquire language, learners must represent both local and non-local dependencies. A substantial body of research has shown that infants are attuned to a range of local dependencies within their first year of life (Dye et al., 2019). These early sensitivities provide a foundation for parsing sentences and uncovering grammatical rules. However, how learners go beyond this foundation and generalize to non-local dependencies is less clear. A key question is when and how infants can represent and interpret more complex structures.

A growing body of research reveals that sensitivity to morphosyntactic dependencies emerges from early on. In a seminal study using the Headturn Preference Procedure, Santelmann and Jusczyk (1998) found that 18-month-olds, but not 15-month-olds, could distinguish between grammatical passages (e.g., *everybody is baking bread*), with a matched auxiliary *is*, and ungrammatical ones, with an unmatched modal auxiliary *can* (e.g., *everybody can baking bread*). This suggests that sensitivity to such dependencies may emerge only after children have observed enough pieces from natural input to detect both the surface pattern and its structural implications. Interestingly, when exposed to more varied and intense input in lab conditions, even younger infants can succeed (Friederici et al., 2011; Omaki et al., 2015). For instance, ERP studies reveal that even 4-month-old German-learning infants can track similar morpho-syntactic dependencies in a non-native language like Italian (e.g., *sta cantando 'is singing'*) after sufficient exposure during familiarization (Friederici et al., 2011). Consistent with this, artificial language studies (R. L. Gómez, 2002; R. Gómez and Maye, 2005) demonstrate that infants succeed in learning these dependencies in labs only when the intervening material is highly variable—specifically, with 18 to 24 types of intervening elements between dependent morphemes.

Movement dependencies, like those in *wh*-questions, share the property of non-adjacency with morpho-syntactic dependencies but differ crucially in one aspect: the second element in the dependency is silent—a gap. The gap itself, being an empty category, may serve as a noticeable local signal, particularly when it follows a transitive category that learners already recognize as requiring a complement.

Consider the object *wh*-question *What did Bowie break?* The verb *break* is transitive and selects for a direct object. When *what* is fronted, the verb’s object position is left empty, creating a syntactic gap. If the learner already knows that *break* typically occurs with a patient argument, the absence of that local argument may serve as a cue that something is missing. It can have both syntactic and semantic consequences—the transitive verb *break* lacks a direct object, and the patient of the breaking event is unidentified. In this way, learners can interpret or respond to *wh*-questions based on their verb knowledge, even without fully representing the movement dependencies involved (Gagliardi et al., 2016; Perkins and Lidz, 2020). That is, they can recover the information that these argument questions are seeking by detecting the relevant argument gaps. This means that learning verb-specific selectional properties from the input might be crucial for learners to detect the direct object gaps when present.

In contrast, identifying morpho-syntactic dependencies, as with *is* and *-ing*, requires learners to attend to both ends of the dependency from the start. As the verb stem preceding the second element *-ing* can surface with or without the *-ing* suffix, there are no reliable local cues signaling that a structural dependency is present. This lack of overt local signals may explain why infants require more varied input to detect such patterns. They need to observe enough instances of the relevant auxiliary-verb pairings, with diverse intervening verbs, to detect the underlying dependency.

This distinction in surface signals between morpho-syntactic and movement dependencies is reflected in developmental patterns. While older toddlers often outform

younger ones in detecting morpho-syntactic dependencies (Santelmann and Jusczyk, 1998), the opposite pattern has been observed for movement dependencies. Gagliardi, Mease, and Lidz (2016) found that 15-month-olds, but not 20-month-olds, successfully interpreted both *wh*-questions and relative clauses involving filler-gap dependencies. They proposed that 15-month-olds rely on a interpretive heuristic based on their knowledge of argument structure. For instance, when hearing an object *wh*-question, like *which dog did the cat bump*, or an object relative-clause, like *show me the dog that the cat bumped*, infants may have responded correctly through recognizing a locally missing direct object of a transitive verb *bump* and then finding the event participant, i.e., the patient, that corresponds conceptually to the missing direct object.

Supporting this view, Perkins and Lidz (2020) further showed that infants’ performance in these tasks correlates with their expressive vocabulary size, a proxy for their knowledge of verbs. Infants with a higher vocabulary were better able to identify and resolve gaps. This led them to conclude that young learners by 15 months may interpret *wh*-questions not by representing syntactic movement per se, but by leveraging their knowledge of local argument structure to “fill in” the missing element. In other words, they use partial representations to approximate adult-like interpretations, effectively “cheating” at these questions without fully parsing them.

A broader proposal emerges from these findings (Gagliardi et al., 2016; Perkins and Lidz, 2020): knowledge of local dependencies—especially verb-based argument structure—may bootstrap learners’ understanding of non-local movement dependencies. Infants recognize argument gaps after transitive verbs in questions and use this cue to infer the underlying intent for them to search for an unmentioned event participant that thematically matches the missing argument (Gagliardi et al., 2016; Perkins and Lidz, 2020).

However, this raises an important question: Is the gap-filling strategy specific to

verb argument structure, or can it extend to other types of argument gaps—such as those introduced by prepositions? If the strategy is general, learners should be able to detect missing arguments in other syntactic contexts where certain categories like prepositions select for obligatory complements.

There is some evidence that learners around 16 months understand the argument-taking properties of the preposition *with*. Lidz, White, and Baier (2017) showed that 16-month-olds treat the preposition *with* as introducing arguments with instrument roles. In their study, infants were divided into two groups: one heard sentences with *with* (e.g., *She’s pushing with the tig*), while the other heard the same noun introduced by a nonsense preposition (*gub*: *She’s pushing gub the tig*). When later asked *Which one’s the tig?*, only infants in the *with* condition looked more to the instrument than to the patient of the action. This suggests they correctly interpreted *with* as signaling a following argument with an instrument role. Crucially, this correct looking pattern emerged despite infants’ general baseline preference for interpreting novel nouns as naming patients rather than instruments.

Still, even if infants understand that *with* takes an instrument-specifying noun phrase, they may struggle to respond to argument questions involving prepositional gaps for two usage-related reasons. First, prepositional gaps by themselves may be sparse in infant-directed speech. Previous research has reported that among *what*-questions, those featuring direct object gaps (89.1%) are far more frequent than those with prepositional object gaps (10.9%) in child-directed corpora (see Atkinson et al., 2018). This quantitative asymmetry suggests that learners may be far more familiar with, and therefore more successful at interpreting, direct object questions than prepositional object ones. This is very likely, considering how input-related factors like item variability and vocabulary size have been found to constrain young learners’ performance in lab conditions with both types of non-adjacent dependencies

(morpho-syntactic dependencies: Santelmann and Jusczyk, 1998; R. L. Gómez, 2002; movement dependencies: Perkins and Lidz, 2020).

Second, prepositional phrases that introduce arguments with instrument roles are syntactically optional. Instrumental *with*-phrases have been traditionally analyzed as modifiers, not core arguments, of verbs (Goldberg, 2002; Levin and Rappaport, 1988; Jackendoff, 1992; Marantz, 1981). Unlike subjects and direct objects, which are integral to the argument structure of a transitive verb, instrumental *with*-phrases are adjuncts that are syntactically optional (Dowty, 2003). For instance, unlike direct objects (8), adjuncts like *with*-phrases can be omitted via VP ellipsis (7). This optionality may make it harder for learners to anticipate a missing argument in a *with*-phrase.

(7) Joe ate the cake with a spoon and Poe did so with a fork.

(8) *Joe ate the cake and Poe did so the donut.

Moreover, both child and adult studies highlight the role of verbs in incrementally shaping the processing of upcoming prepositional phrases. Lidz et al. (2017) found that 19-month-old English-learning toddlers often misinterpret sentences like *She is pulling with the tig* by treating *the tig* as the direct object of *pull*, and effectively ignoring the preposition *with*. This error is likely attributed to a strong transitivity bias—children expect transitive verbs like *pull* to be followed by a direct object, even when a preposition intervenes. Properties of verbs—whether they typically specify an instrument—are represented in a gradient way (Rissman et al., 2015) and have been found to influence adults’ resolution of prepositional gaps. Koenig, Mauner, and Bienvenue (2003) showed that adults process *with*-questions more easily when the verb strongly implies an instrument (e.g., *which sword did the rebels **behead** the*

traitor king with _____ during the rebellion) than when it does not (e.g., *which sword did the rebels **kill** the traitor king with _____ during the rebellion*). This suggests that verbs entailing an instrument (e.g., *behead*) may facilitate the resolution of a prepositional gap introduced by *with*, making interpretation easier.

Considering the special role of verbs in sentence processing with both toddlers and adults, it is possible that resolving prepositional gaps in *with*-questions may be more difficult for early toddlers, especially with verbs that do not typically occur with an instrument-specifying argument. As few early-acquired verbs entail an instrument in their meaning, young learners are unlikely to expect a prepositional phrase after a verb, let alone an argument gap within that prepositional phrase. As a result, when faced with an interrogative clause containing a final stranded *with*, learners may not promptly look to the corresponding instrument in a comprehension task. They may struggle to identify the missing instrument argument, simply because the verb itself provides little subcategorization cue for such an optional structure.

Taken together, the existing literature suggests that learners at 15 months use knowledge of verb argument structure to respond correctly to *wh*-questions. However, we still lack a clear understanding of how general their gap-filling strategy is in helping them “cheat” at question comprehension. Can it extend to *wh*-questions with prepositional object gaps? Does the recognition of a missing prepositional argument in a question like *What did she break the stick with?* prompt the same type of interpretive response seen with direct object gaps? Would usage-related factors of instrumental *with*-questions cause troubles for interpretation? Before tackling this question, we must consider the cognitive prerequisites that infants might bring to the task, particularly how they conceptualize events. The next section provides a brief overview of when and how infants represent these thematic roles, independent of language.

2.1.2 Prerequisites for linking thematic roles

Understanding how infants represent events is crucial for uncovering the cognitive foundations that allow young learners to navigate the complexities of language. Long before they acquire a full-fledged grammar, infants must track who is doing what to whom in dynamic scenes, identifying not just the participants but also the structure of their interactions. Critically, they must also learn what the surface signatures of various argument positions are—that is, how linguistic form signals underlying roles like agent, patient, or instrument. If infants can readily extract such roles from events, this suggests a deep, early-emerging system of conceptual organization that supports both event cognition and language acquisition. The question, then, is how rich and structured this early event representation is: which thematic roles are infants sensitive to, and how do they integrate these roles in understanding the world?

One of the most fundamental distinctions in event representation is between agents and patients, namely entities that initiate and control actions and those that undergo or are affected by those actions. These roles have been argued to be part of infants' core knowledge, forming an early cognitive framework for understanding causality and action (Carey, 2009). By 5 months of age, infants already attribute goals to some action of intentional agents (Woodward, 1998; see Sommerville et al., 2005 for evidence in 3-month-olds). This means that they recognize that an agent's action is often oriented toward an intended outcome. By around 7 months, infants begin to differentiate specific roles within events (Rochat et al., 2004; Papeo et al., 2024). For instance, they can distinguish between a chaser (the agent) and a chasee (the patient), and detect when these roles are reversed, indicating an early ability to track relational roles in dynamic interactions (Rochat et al., 2004). By 12 months, their event representation extends beyond directly perceivable agents: they can infer

the presence of hidden causal agents responsible for inanimate objects in motion, demonstrating a more sophisticated understanding of agency (Saxe et al., 2005). These findings collectively highlight that infants’ ability to perceive and reason about distinct thematic roles emerges early and becomes increasingly refined within the first year of life.

While agents and patients are central to infants’ early event representations, they are not the only critical participants in an event—many actions depend on the use of instruments. Instruments are the means through which agents achieve their goals, forming a crucial bridge between intentional action and outcome. With respect to their linguistic importance, Fillmore (1968) also proposed instruments as part of a small set of universal thematic roles.

If instruments are as fundamental as agents and patients, we should expect to see early-developing sensitivities to their presence and function in infants. Empirical evidence supports this expectation. By as early as 4 months, infants begin to individuate and track simple tools, provided they have observed contrastive functional demonstrations that highlight the tool’s role in achieving a particular goal (Stavans and Baillargeon, 2018). By 8 months, infants can infer the causal power of an intentional agent using a tool to change the state of an inanimate object (Adibpour and Hochmann, 2023). Notably, they expect that a tool should only bring about an effect through contact with the patient if it was wielded by an intentional agent. This suggests that they integrate both agency and means-end reasoning in their event representations. By 11 months, infants perceive and register instruments as an explicit participant of an event (He, 2015). In one study by He in her dissertation, after habituating to a scene of a girl opening a box, infants showed greater dishabituation when the scene switched to one where she used a lever to jimmy the box open than when she opened the box from a different orientation. This suggests that infants do

not just perceive instruments as incidental objects but as central components of an event structure. By 12 months, infants’ ability to process event sequences becomes more hierarchical. They interpret multi-step actions with respect to an ultimate goal, recognizing that the use of an instrument is often an intermediary step rather than the final outcome itself (Sommerville and Woodward, 2005).

In sum, these findings reveal a striking convergence in the developmental timelines of agent-patient and instrument representations. Infants possess a richly structured cognitive system for analyzing events, one that organizes thematic roles within an intentional framework—tracking who is acting, what is being acted upon, and what means are being used to carry out the action. This system not only enables them to make sense of the physical and social world but also likely scaffolds their acquisition of argument structure in language (Pinker, 1984; Williams, 2015; Perkins et al., 2024).

Although infants appear to have early access to these thematic roles, not all of these roles are equally well represented in surface syntax. Unlike subjects and direct objects, which often surface in sentences with transitive verbs, prepositional objects—such as instruments introduced by the preposition *with*—are syntactically optional and occur much less frequently. Stranded *with*-phrases, where the complement of *with* is displaced (as in *What did she break the stick with?*), are potentially even rarer in early input. This disparity raises an important question: if infants already understand the concept of instruments and represent them conceptually as participants of events, and that by around 16 months, they know *with*-phrases can introduce them, can they apply this knowledge to interpret instrumental *with*-questions, when the complement of *with* is missing? Or will they fail to resolve prepositional object gaps due to usage factors?

In Study 1, I present a corpus study demonstrating that instrumental *with*-questions are indeed sparse in infant-directed speech to 9- to 15-month-olds. How-

ever, I also show that learners are exposed to a broader set of prepositional object questions involving different prepositions, which may provide a foundation for generalizing to instrumental *with*-questions. In Study 2, I report that by 15 months, English-learning children successfully respond to instrumental *with*-questions. This supports the transitivity-driven hypothesis, which predicts learners at this age to correctly identify the missing argument signaled by the highly transitive preposition *with*. These findings suggest that early toddlers’ comprehension of argument questions is not narrowly tied to usage-related factors like the frequency of the question type or verb-specific expectations. Rather, it reflects a broader sensitivity to transitivity—a category-general property that allows learners to recognize the presence of syntactic gaps, regardless of whether they are introduced by transitive verbs or prepositions.

2.2 Study 1

In this study, we aim to first examine the types of *what*-questions 15-month-olds receive in their input. The transitivity-driven hypothesis relies on the assumption that learners can identify the gaps following transitive categories. As prior literature (Atkinson et al., 2018) focuses on input of *what*-questions to children within a broad age range, this corpus study will help us assess how prepositions behave in early input with respect to their transitivity. What generalizations may learners arrive at about *what*-questions in general? How can these generalizations inform learners about instrumental *with*-questions?

2.2.1 Method

The analysis is based on the Brent corpus on CHILDES (Brent and Siskind, 2001). The corpus includes conversations from 16 mother-child pairs when each child was

between 9 and 15 months of age. The mothers were all English-speaking—half of them had higher education (at least a bachelor’s degree) and half of them lower education (none completed more than 2 years of college). The corpus consists of a total of 364,220 words and 123,579 sentences.

I first focus on *what*-questions produced by mothers, and label the questions as subject (S) questions, direct object (DO) questions, and prepositional object (PO) questions, as has been done in previous research (Atkinson et al., 2018). I exclude *what*-questions that contain a *what*-phrase (e.g., *what color*), serve as exclamatives (e.g., *what!*), or are identified as labeling or copula questions (e.g., *what is that*).

Additionally, I identify the verbs occurring in these DO questions and the prepositions occurring in these PO questions. Based on this list of gap-introducing verbs and prepositions in the corpora, I further calculate their occurrences in declaratives and derive their likelihood of undergoing object extraction, i.e., introducing either DO gaps or PO gaps.

2.2.2 Results

Our corpus analysis reveals a robust asymmetry between direct object (DO) and prepositional object (PO) *what*-questions in early child-directed speech. As shown in Figure 2.1 for *what*-questions, PO *what*-questions are far less frequent compared to DO questions. This replicates earlier findings reporting a similar disparity between DO gaps and PO gaps in child corpora covering a broader age range (1;1-5;1, see Atkinson et al., 2018). For additional breakdowns by caregiver education, see Table A.1 in Appendix A. For differences by corpus source, see Figure A.1 and Table A.2 in Appendix A.

To contextualize this disparity, we examined the base rate of transitive verbs to

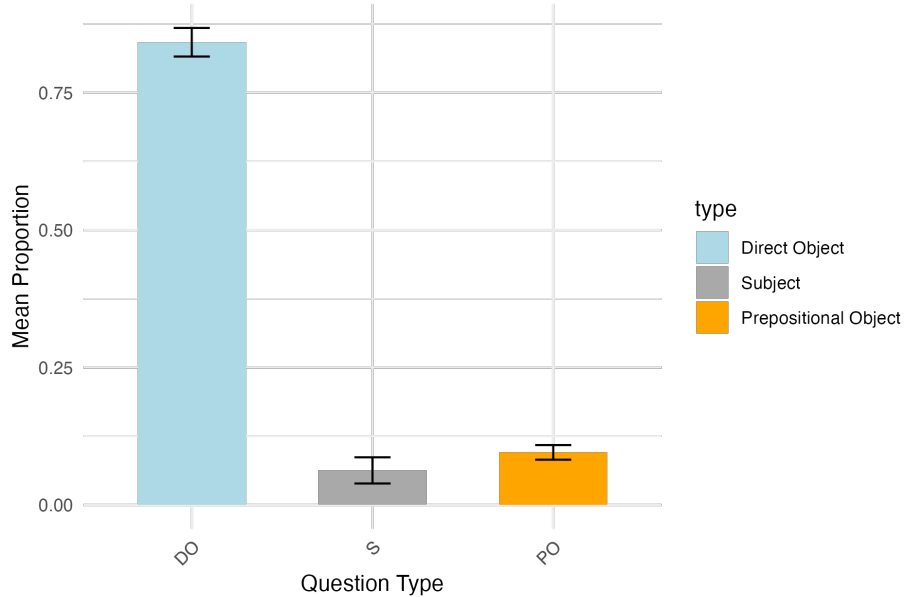


Figure 2.1: Proportion of Question Type by Corpus

transitive prepositions in declarative clauses. If the rate of transitive verbs to transitive prepositions in *what*-questions is even higher than their baseline rate in declaratives, this would suggest that transitive prepositions are even more underrepresented in *what*-questions, compared to the baseline in declaratives. Indeed, this is borne out in our corpora. As indicated in Table 2.1, in declaratives, where no syntactic extractions take place, the ratio of transitive verbs to transitive prepositions is 1.43. However, in *what*-questions where syntactic extractions occur, this ratio skyrockets to 10.44. This shift indicates that transitive verbs are substantially more likely than transitive prepositions to appear in gap-introducing environments. In other words, while both transitive verbs and prepositions appear in declaratives, verbs disproportionately dominate in the formation of argument gaps in *what*-questions.

To investigate which prepositions are most likely to introduce a syntactic gap in PO *what*-questions, we calculated the average proportion of PO gaps by preposition, relative to their baseline use in declaratives. Figure 2.2 plots these results. Across

	Verb	Preposition	Verb/Preposition
Declaratives	9035	6300	1.43
Questions	1399	134	10.44
%gap	13.41	2.08	

Table 2.1: Verb-preposition rate in declaratives vs. *what*-questions

the board, the average rate of gap formation is low, typically hovering around 2%.

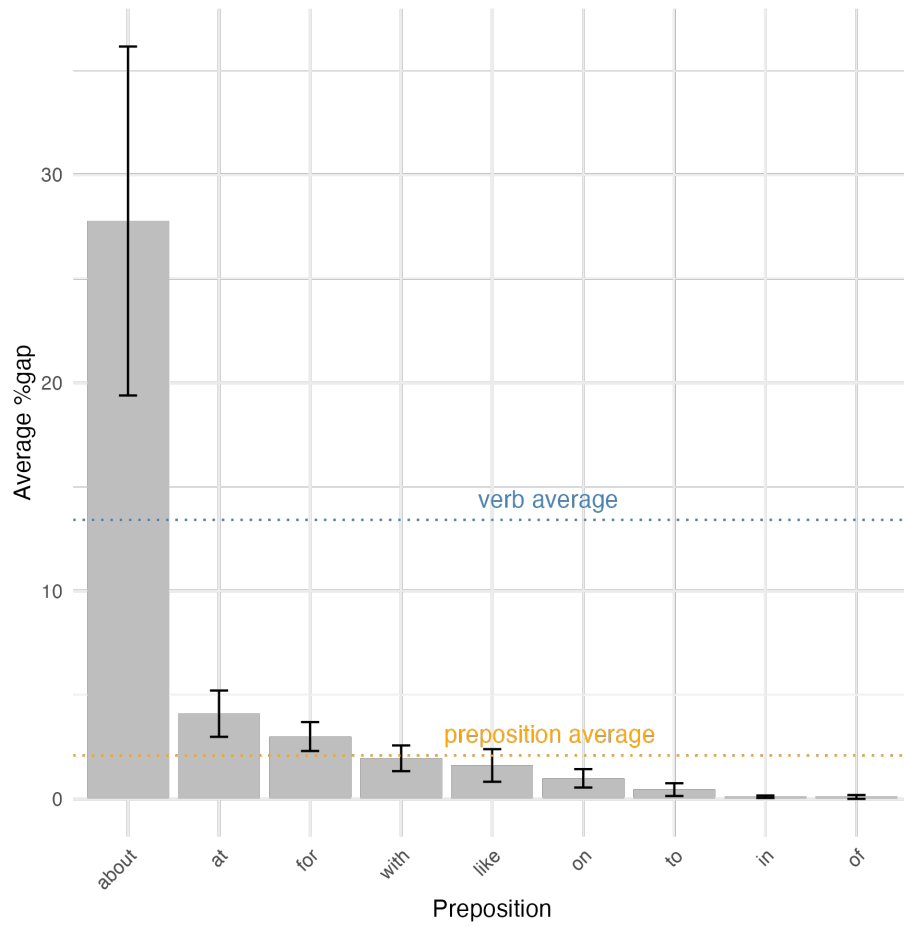


Figure 2.2: Average proportion of gap by preposition

A few prepositions, however, stand out in PO *what*-questions: *about* (27.8%), *at* (4.1%), and *for* (3.0%). These are the prepositions that most frequently introduce PO gaps across corpora. The examples below show instances of these questions in

the corpora.

(9) What are you fussing about now?

(v2, 00;09;03)

(10) What is Dakota barking at outside?

(s2, 00;10;18)

(11) What are you looking for?

(v1, 00;10;24)

At the individual corpus level, both the frequency and the proportion of PO gaps remain generally low. The heatmap in Figure 2.3 visualizes how often each preposition appears in PO *what*-questions across different corpora. The y-axis ranks how common it is to observe a particular preposition introducing prepositional gaps in *what*-questions across corpora, with the most common preposition at the top. Similarly, Figure 2.4 presents proportional data, confirming that PO gaps remain rare regardless of corpus source.

Notably, as can be seen in Figure 2.3 and Figure 2.4, *with*-questions are especially sparse, showing up in only half of the corpora (for the proportion of *with*-gaps in different corpora, see Table A.3 in Appendix A). When collapsed across corpora, these questions make up just 1.1% of all DO and PO *what*-questions. Crucially, none of these *with*-questions involve an instrumental interpretation. Instead, the preposition *with* in these questions consistently reflect a comitative reading, as illustrated in the examples below.

(12) What haven't we played with in a while?

(d1, 00;09;04)

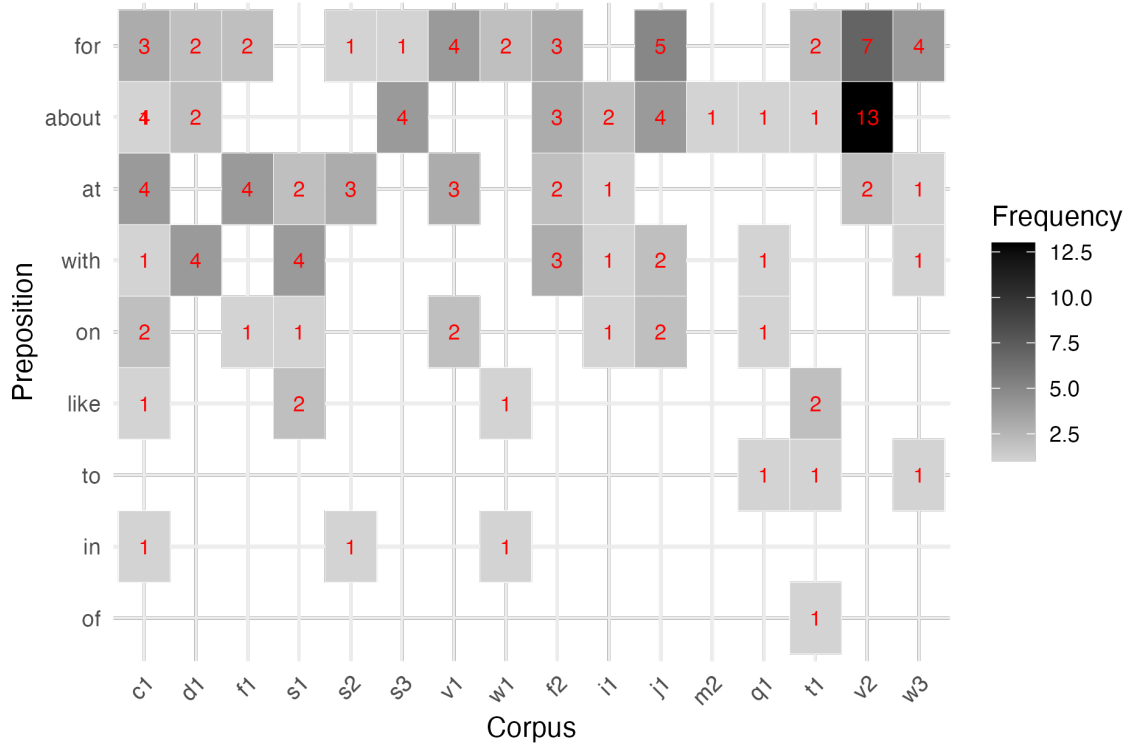


Figure 2.3: Heatmap of preposition frequency by Corpus (occurrences marked within cells)

(13) What do you sleep with, Moomoo?

(j1, 00;10;23)

(14) What are you meddling with?

(w3, 01;01;01)

We further assessed the availability of instrumental readings of *with* outside of questions, namely in declaratives. Out of 702 *with*-phrases found in declaratives, only 42 (6%) were used instrumentally. These were distributed across only 16 verb types, representing a mere 1.8% of the total 878 verb types in the corpora. This indicates that instrumental *with*-phrases are not only rare, but also highly restricted in the verbs they surface with. The following examples illustrate instrumental uses of *with* in declaratives.

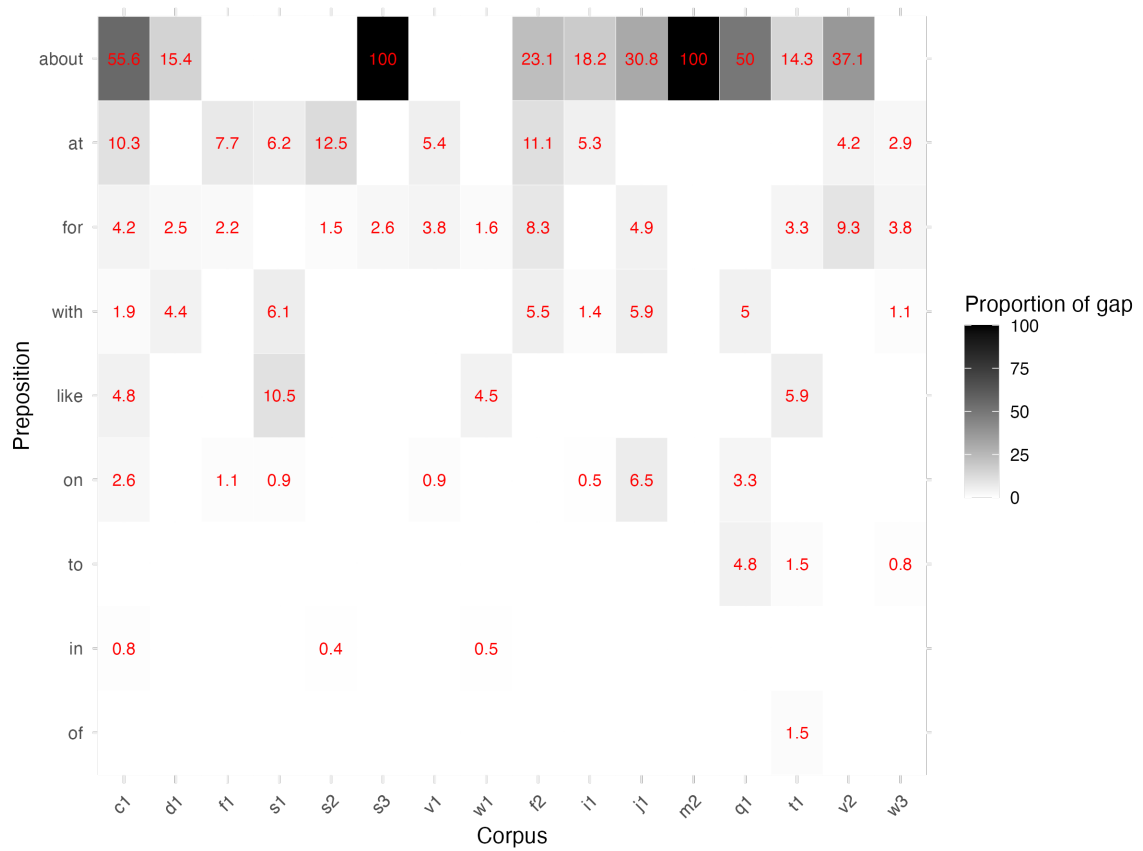


Figure 2.4: Proportion of gap by preposition across corpora (proportion of gap marked within cells)

(15) You can hit it with the mallet.

(i1, 01;00;27)

(16) You eat it all up with the spoon.

(w3, 01;00;08)

(17) Mommy will pick this up with her foot.

(t1, 00;10;25)

We also calculated the total occurrences of these instrument-compatible verbs in the corpus to further examine the likelihood of them occurring with instrumental

with-phrases. Table 2.2 lists these verbs and the frequency of their co-occurrence with instrumental *with*-phrases in declaratives. As seen from the table, even among this restricted set, co-occurrence rates are extremely low. Only a handful of verbs—e.g., *hit*, *whip*, *tickle*, *mark*—are slightly more likely to occur with instrumental *with*-phrases, and together, they account for just 0.5% of all verb types. These results underscore that in early child-directed input, instrumental *with*-phrases are not only infrequent, but also narrowly distributed across an extremely small set of verbs.

verb	instrumental use	%instrumental	other uses	%other	total
touch	1	2.3%	43	97.7%	44
hold	5	5.7%	82	94.3%	87
eat	5	1.0%	499	99.0%	504
hit	10	13.9%	62	86.1%	72
reach	1	2.3%	42	97.7%	43
wash	2	1.0%	198	99%	200
comb	1	4.3%	22	95.7%	23
pick	1	1.4%	71	98.6%	72
rub	1	7.7%	12	92.3%	13
hurt	1	0.8%	119	99.2%	120
bite	1	1.1%	87	98.9%	88
whip	1	33.3%	2	66.7%	3
play	1	0.2%	462	99.8%	463
tickle	7	12.7%	48	87.3%	55
mark	3	75.0%	1	25.0%	4
wake	1	4.0%	24	96.0%	25
total	42	2.3%	1774	97.7%	1816

Table 2.2: Proportion of instrumental *with*-specification by verb

Taken together, these findings highlight the extreme sparsity of instrumental *with*-questions in the input to learners between 9 and 15 months. Even in declarative clauses—where no gap is required—the majority of verbs do not appear with instru-

mental *with*-phrases. This has direct implications for theories of gap resolution.

If these young learners adopt a transitivity-driven strategy, focusing on the selectional properties of the gap-introducing category, they should succeed. Locally, the preposition *with* is highly transitive across corpora (see Table A.3 in Appendix A). This broader structural regularity, supporting an anticipation of following arguments, might enable 15-month-olds to detect a missing argument following *with*. However, they may fail at resolving these instrumental *with*-questions, if their performance is contingent on usage-related factors like the frequency of the question type or verb-specific expectations.

2.3 Study 2

In Study 1, I examined corpora of infant-directed speech to 9- to 15-month-olds and established that prepositional object gaps are much less frequent than direct object gaps in the input, and that instrumental *with*-questions are even more sparse.

In this light, 15-month-olds may fail to comprehend instrumental *with*-questions, if their responses to questions are contingent on usage-related factors. However, they should succeed if their responses to questions are generally driven by the detection of locally missing arguments, whether after verbs or prepositions. As the preposition *with* is almost always followed by an argument, 15-month-olds should respond correctly to prepositional object gaps in *with*-questions.

2.3.1 Method

Study 2 tested whether 15-month-old English infants respond correctly to instrumental *with*-questions like *What did she break the stick with*. Infants were familiarized to a BREAKING scene: a girl tries to break a stick but fails, and then she uses a ham-

mer and succeeds in breaking the stick, with audio describing the scene. After these familiarization trials, infants should represent three participants for such events—the agent (girl), the patient (stick), and the instrument (hammer).

At test, infants were presented with two pictures side-by-side and asked an instrumental *with*-question like *What did she break the stick with*. One picture showed the patient (the broken stick), and the other the instrument (the hammer). By measuring infants’ looking preference towards either the patient or the instrument, we can assess whether they can correctly respond to the instrument question being asked. They should look more to the instrument if their response to argument questions reflects a search for an unmentioned event participant corresponding to the missing argument after the preposition *with*. However, if their response to questions is not guided by the signal from the missing argument in a question, their tendency to look at the instrument may be reduced, in view of a potential baseline preference for looking at the patient, as observed in prior studies (Lidz et al., 2017).

Participants. Participants included 24 typically-developing infants (13 males) between the ages of 1 year 2 months 16 days and 1 year 4 months 17 days (mean = 1 years 3 months 19 days). Participants were recruited from the greater Washington, D.C. area. Parents provided information on their child’s total productive vocabulary using the Words and Sentence MacArthur-Bates Communicative Development Inventory (MCDI; Fenson et al., 1993). Mean total words produced were 16.3 (range: 0-55), with a median of 11.

Materials. Visual stimuli featured video recordings of an actress performing various actions with inanimate objects. Multiple tokens of each event were filmed and later edited in Wondershare Filmora 11 to create videos for the trial structure in Table 2.3.

Audio stimuli were recorded by a female native speaker of American English using

Phase	Audio	Video
Event	<i>Look! A girl. She wants to break the stick! Oh no! She can't break it. Now she's got something. Let's see! Yay! She used it to break the stick!</i>	Girl breaks stick with hammer (25s)
Test	Silence	Black screen (1s)
	<i>Did you see that? What happened?</i>	Both objects (2s)
Baseline	Silence	Both objects (2s)
	<i>What did she break the stick with?</i>	Both objects (2s)
Test1	Silence	Both objects (3s)
	<i>What did she break the stick with?</i>	Black screen (2s)
Test2	Silence	Both objects (3s)

Table 2.3: Structure of Experimental Trial, Study 1

child-directed speech. Stimuli were edited and synchronized with the video stimuli using Wondershare Filmora 11.

Procedure. Infants were seated either on a parent’s lap or in a high chair at home, facing a laptop screen. Parents were instructed to remain silent and not direct their child’s attention. Stimuli were presented using Slides.com, and a laptop camera recorded the session. A zoom host guided parents in adjusting the laptop position or their distance from the camera to ensure optimal video quality for offline coding.

Each experimental session lasted 4.7 minutes. The structure is as follows. Infants first watched a 7-second video of flickering dots, accompanied with a short clip of music from *The Four Seasons* by Antonio Vivaldi. This was followed by 6 trials, each consisting of an event phase and a test phase (Table 2.3). During the event phase, infants viewed a 25-second video depicting an event scene. For instance, a BREAKING scene was described by the following.

(18) *Look! There’s a girl. She wants to break the stick. Oh no! She can’t break it.*

Now she's got something. Let's see! Yay! She used it to break the stick.

Figure 2.5 shows sample snapshots from the test videos in the experimental trial (see Appendix B for the full set of stimuli).



Figure 2.5: Images of an experimental video for a BREAKING scene



Figure 2.6: Image of a test phase for the BREAKING scene

During the test phase, two images were presented on different sides of the screen: the patient (e.g., the broken stick) and the instrument (e.g., the hammer) from the previous scene (Figure 2.6). The positioning of the patient and the instrument on the screen was counterbalanced across participants. Additional pseudo-randomization ensured that the sides of the patient and the instrument varied across trials. Each test

phase began with a neutral question prompt (*Did you see that? What happened?*), followed by a 2-second silence. This allowed infants to inspect both images before hearing the test sentence. This step controlled for potential baseline preferences for one of the two images due to visual salience. A test sentence (e.g., *What did she break the stick with?*) or a control sentence (*Did you see that?*) was then played. The 3 seconds following the first test sentence served as the first test window of interest (Test1). Next, the screen briefly turned black, while the same test or control sentence was repeated. This was to ensure that infants had time to process the question without visual distractions. After the repeated question, the two images reappeared in their original positions. This marks the start of the second test window of interest (Test2), which also lasted 3 seconds.

To assess infants' comprehension of instrumental *with*-questions, the study included six test trials, each centered around a different change-of-state verb. These verbs—*open*, *close*, *break*, *clean*, *fix*, and *turn*—were chosen because they clearly depict a change in the state of an object, making the outcome of the event visually salient and thus easier for infants to track.

All six verbs are highly frequent in child-directed speech, as reported in previous research (Perkins et al., 2022). Importantly, each selected verb is highly transitive, meaning that it typically occurs with a direct object. The proportions of these verbs taking a direct object are summarized in Table 2.4, based on speech to 9- to 15-month-olds (Brent corpus on CHILDES).

Crucially, none of the chosen verbs occur with *with*-phrases in declarative clauses in infant-directed speech. This design was deliberate: by selecting verbs that do not typically occur with *with*-phrases, we controlled for potential lexical biases. As a result, infants could not rely on familiar verb-driven expectations to anticipate what comes after the verb phrase in the test sentences. This helps ensure that any

Verb	Total	#DO	%DO
open	289	231	79.9%
close	177	167	94.4%
break	37	27	73.0%
clean	175	96	54.9%
fix	64	61	95.3%
turn	369	330	89.4%

Table 2.4: Test verbs: Counts and percentage uses with direct objects in the Brent corpus on CHILDES

success in interpreting the *with*-questions reflects sensitivity to argument structure and transitivity, rather than rote familiarity with specific verb-preposition pairings.

The six corresponding events presented to the participants were: opening a jar, closing a box, breaking a stick, cleaning a board, fixing a plate, and turning a toy truck. To maintain infants’ engagement throughout the task, the trials were interleaved with brief attention-getting videos. A longer, 7-second attention-getter was placed between the third and the fourth trial to provide a short break, while the others were 5 seconds long.

Predictions. If 15-month-old infants use a general gap-filling strategy to respond to questions, namely via detecting missing arguments given the selectional requirements of transitive categories, this predicts that they should succeed in responding to *with*-questions. They should notice the missing argument after the preposition *with* at the end of the question and look to the instrument that fulfills the role of the missing argument.

However, if infants’ response to questions are influenced by usage-related factors, like the frequency of question type or verb-specific expectations, they are more likely to fail at comprehending *with*-questions—not showing above-chance preferential looking for the instrument. As shown in Study 1, compared to transitive verbs,

prepositions are far less likely to miss a local argument in interrogatives, against their baseline rate in declaratives. Moreover, instrumental *with*-questions are by themselves sparse in input, which might add to the difficulty of resolving the gap following the preposition *with* in particular. Furthermore, most verbs in early input do not typically cooccur with *with*-phrases that specify instruments, even in declarative clauses.

2.3.2 Results

Data Preparation. The recorded sessions of the test phase for each experimental trial were coded frame-by-frame from the onset of test images on the screen. Data were also coded by two other experimenters, with an intercoder reliability above 90% (Cohen’s Kappa > 0.90).

Analysis. Frame-by-frame analysis calculated for each frame, whether each participant was looking at the patient image, the instrument image, or neither. These looking preferences were then averaged across participants to create a time course of the proportion of target looks at the instrument image at each frame, out of looks at either image. That is, the dependent measure is the proportion of looks toward the instrument divided by the sum of looks toward either the patient (distractor or nontarget) or the instrument (target), as has been practiced as a standard in the infant literature (Waxman et al., 2009).

To assess whether participants’ looking patterns were conditioned on the test sentence, three windows of interest were selected within the time course. The Baseline window spans the 2 seconds after the offset of the baseline question (*Did you see that? What happened*), from 2000 to 4000 msec. The Test1 window spans the 3 seconds after the offset of the first question (Wh condition: *What did she break the stick with*; Control condition: *Did you see that*), from 6000 to 9000 msec. The Test2

window spans the 3 seconds after the offset of the second question (repeating the first question), from 11000 to 14000 msec. Figure 2.7 plots the proportions of target looks to the instrument across the three windows, using the eyetrackingR package (Dink and Ferguson, 2015).

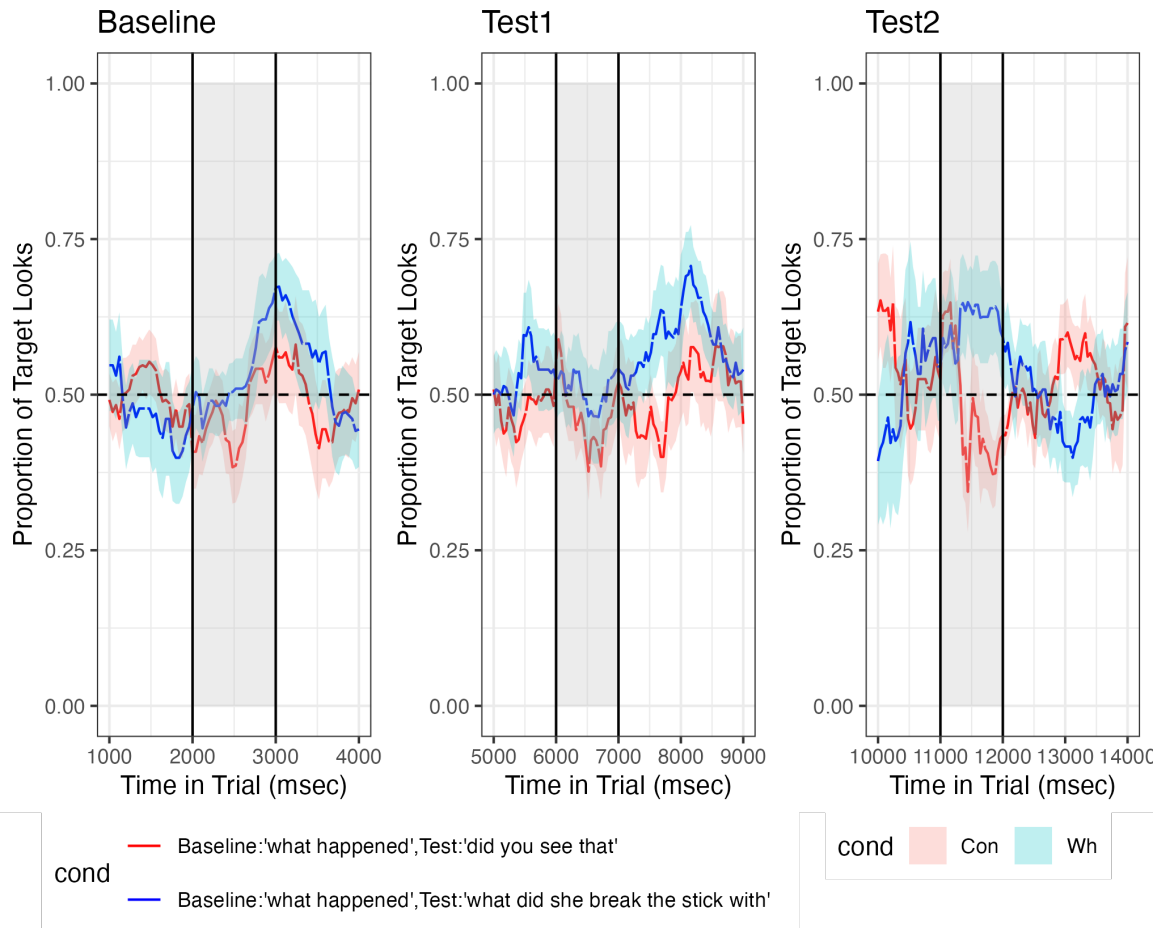


Figure 2.7: Infants' proportion of target looks by delimited Window and Condition (the first vertical line in each facet denotes the offset of the auditory stimuli)

We focus on the first 1000 msec after the offset of the auditory stimuli as the window of analysis, based on Gagliardi et al. (2016). That is, instead of averaging looks to target across a whole trial, we choose a specific region, namely the first 1000 msec following each auditory stimuli. This is because toddlers at this age have been

found in prior research (Gagliardi et al., 2016) to reliably look at the target during the 1000 ms window during Test2, instantly following the repetition of the test question. These delimited windows of analysis are illustrated by the grey regions in Figure 2.7. A 2×3 mixed ANOVA (condition*window) revealed a significant effect of condition, $F(1,22) = 5.13$, $p = .034$, partial $\eta^2 = .078$, but no effect of window or any interaction between condition and window.

First, we observed a distinction in proportion of target looks between the two conditions only in Test2, $t(19) = 2.56$, $p = .019$, namely after the repetition of the instrumental *with*-question. This distinction was not found for either Baseline, $t(22) = .95$, $p = .354$, or Test1, $t(19) = .62$, $p = .544$. Figure 2.8 illustrates the results.

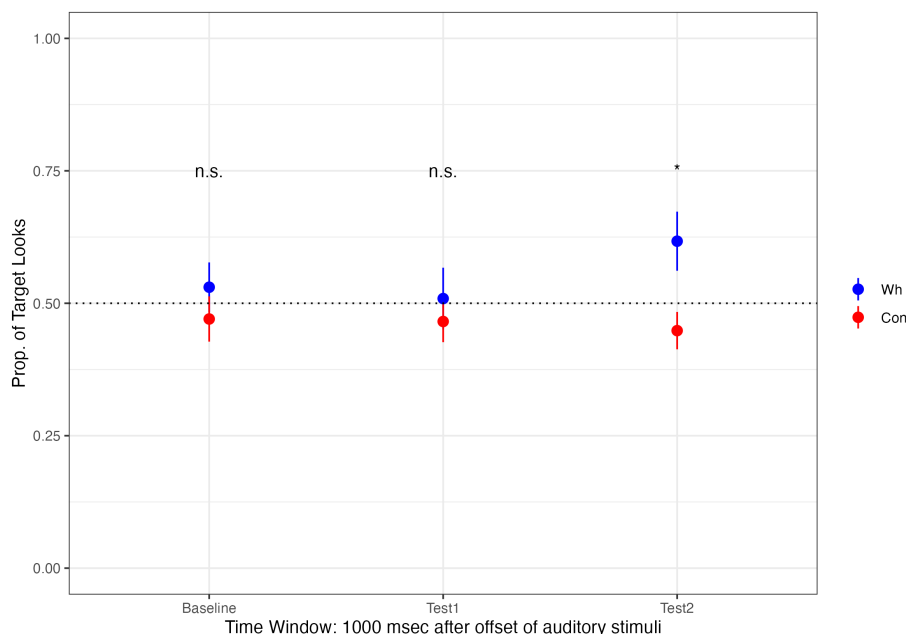


Figure 2.8: Average proportion of target looks to the instrument in the Wh condition compared to the Control condition

These results suggest that 15-month-olds looked significantly longer to the target instrument in the first second after the offset of the question only after the question was repeated. This replicates findings from Gagliardi et al. (2016), where 15-month-

olds also showed looking behavior indicating successful comprehension in the first second after the repetition of the direct object question.

Second, we also found significant looking towards the target instrument above chance level only for 15-month-olds in the Wh condition during Test2, $t(11) = 2.10$, $p = .03$, $M = .62$. This, however, was not observed for Baseline and Test1 in the Wh condition, or for all windows in the Control condition. Figure 2.9 illustrates the results.

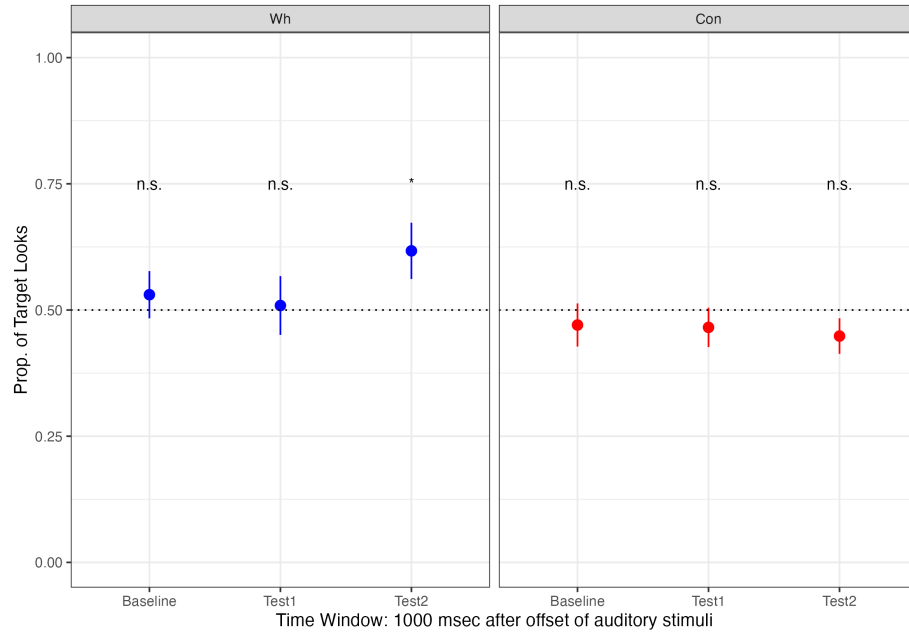


Figure 2.9: Average proportion of target looks to the instrument across three delimited windows in the Wh vs. Control condition

These findings suggest that 15-month-olds showed successful interpretation in the first second only for the Test2, but not Test1 window, compared to the baseline window. No such preference for the target instrument was observed during any window in the Control condition (see Table A.4 in Appendix A). Rather, in the first second after hearing *Did you see that* in the Control condition, 15-month-olds seem to have a tendency to look towards the patient, which replicates the patient bias observed in

prior research (Lidz et al., 2017).

While the findings so far suggest that 15-month-olds only show successful interpretation in Test2 in the first second following the offset of the question, it is possible that they might show signs of comprehension in a later window the first time they encounter the question. Specifically, by eyeballing the looking data in Figure 2.7, we can observe some sign of increased looking to the target during the second half of the second second (7500 to 8000 msec) in the Test1 window.

Indeed, when we examine sliding epochs of 200 milliseconds with 100 millisecond intervals from 7500 msec to 8000 msec, we found significantly increased target looking in the second epoch, from 1600 msec to 1800 msec after the offset of the first instrument question, $t(21) = -2.18$, $p = .041$, as shown in Figure 2.10.

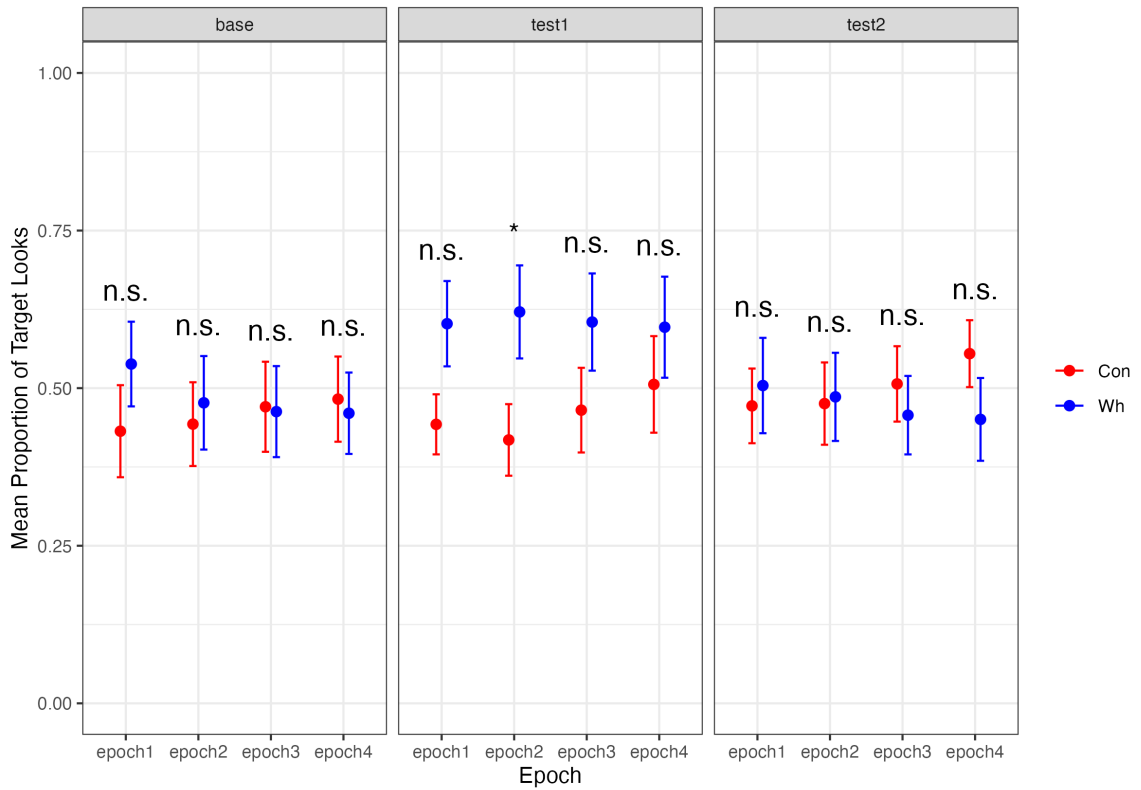


Figure 2.10: 200 millisecond epochs from 1500 msec to 2000 msec after the offset of the questions in different windows

This distinction in looking behavior between the two conditions was observed in no other epochs in the second second following the first question in Test1 and in none of the epochs for the Baseline window or the Test2 window. These results indicate that while 15-month-olds do not show signs of comprehension instantly after they hear the instrument question for the first time (Test1), namely in the first second window, they do show emerging target looking compared to the Control condition later on, towards the second half of the second second.

In sum, infants did not show baseline preference for either the patient or the instrument image during the Baseline window, but they showed a significant preference for the target instrument within the first second after the offset of a repeated *with*-question compared to those hearing the control question. Crucially, this distinct preference for the target instrument above chance level was not observed in the first second window after the offset of the Baseline question or the first occurrence of the target instrument question. Additionally, we found emerging comprehension in a later window following the first occurrence of the target question (Test1), namely in a 200 millisecond narrow window within the second second.

2.4 General discussion

The findings presented in this chapter show that 15-month-olds can correctly interpret instrumental *with*-questions, such as *What did she break the stick with?* This ability supports the transitivity-driven hypothesis, which suggests that 15-month-olds' comprehension of *wh*-questions is guided by a broader sensitivity to missing arguments signaled by transitive categories—those that typically require a noun phrase (NP) complement. In the case of instrumental *with*-questions, infants recognize an argument gap after the preposition *with*—a preposition that selects for an NP bearing

an instrument role—and then search for a corresponding event participant that fulfills the role of the missing argument. Crucially, our data indicate that 15-month-olds’ success in interpreting argument questions is not contingent on usage-related factors such as experience with frequent question types or verb-specific lexical expectations.

As shown in Study 1, questions involving prepositional object gaps are much less frequent than those involving direct object ones. Instrumental *with*-questions, in particular, are extremely sparse in their input. Despite this, infants in our study show correct response to instrument questions, suggesting that infants are capable of resolving missing arguments even in the absence of exposure to such clauses. This is consistent with findings from prior research (Gagliardi et al., 2016; Perkins and Lidz, 2020), where 15-month-olds can interpret not just *wh*-questions available in their input but also relative clauses that are much less frequent.

Moreover, in resolving argument questions, 15-month-olds’ expectation of arguments in clauses is not constrained by subcategorization frame frequencies associated with verbs. Even though the verbs in our study never occur with *with*-phrases in declaratives in infant-directed speech, infants at this age still demonstrate robust target looking response shortly after hearing the repeated *with*-question. This pattern mirrors earlier findings on their performance with direct object questions (Gagliardi et al., 2016) and cannot be explained by verb-driven lexical expectations. If infants at this age show incremental prediction, making parsing commitments to the upcoming syntactic elements based on the verb, they would not anticipate a *with*-phrase following the verb phrase, so when they hear the preposition *with*, they would have to revise their parsing decisions. As revising initial parsing commitments has been shown to be difficult even for preschoolers (Trueswell et al., 1999), we would expect 15-month-olds to fail at comprehending these instrumental *with*-questions.

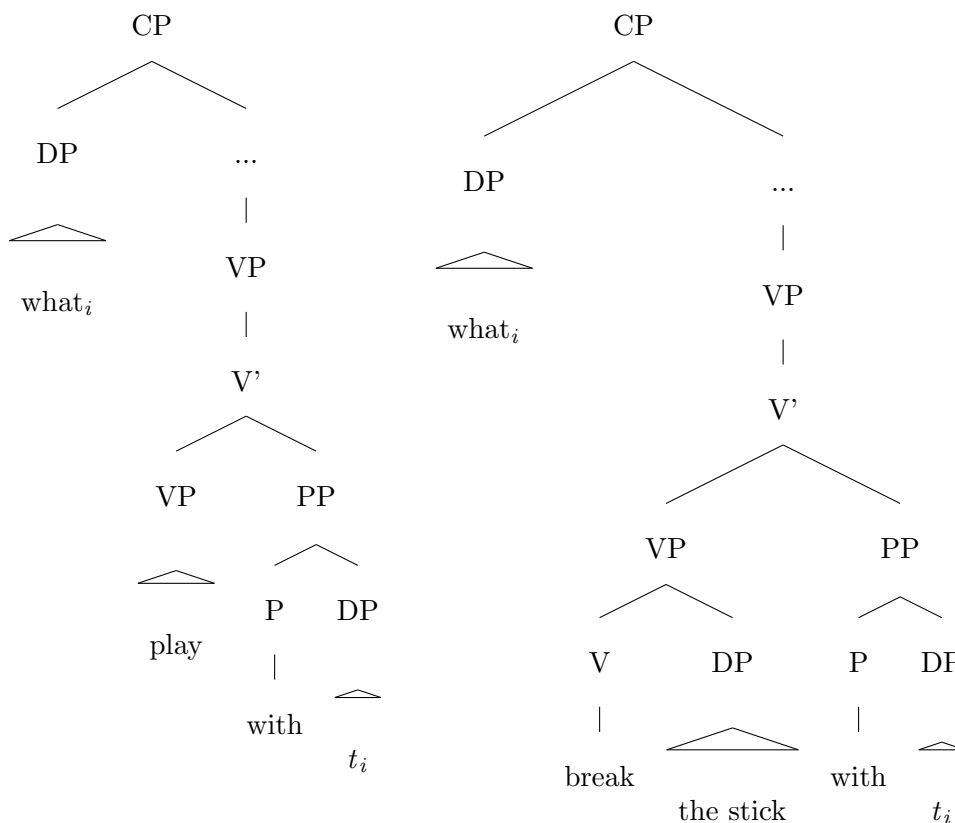
Instead, our results align with those from prior research (Lidz et al., 2017), suggest-

ing that infants at 15 months still process incoming syntactic elements in a veridical, bottom-up fashion. They do not seem to prematurely make strong, lexically based predictions. This may reflect a developmental strategy where early learners are initially guided by bottom-up processing and focus on local structural selections, with predictive processing abilities emerging later (cf. Lidz et al., 2017) as their structural knowledge matures (Perkins and Lidz, 2021).

This ability to interpret rare structures like instrumental *with*-questions also has important implications for how infants generalize from limited input. As Study 1 has illustrated, prepositional object questions are relatively rare, and among them, instrumental *with*-questions are virtually absent in the data. This raises the possibility that these young learners may draw on their limited experience with more common prepositional object questions—those involving prepositions like *for*, *about*, and *at*—to form a broader generalization that English allows gaps following prepositions.¹ By this means, they can generalize beyond what they have heard, such that questions involving missing arguments after the preposition *with* become a possible structure in their language that they might have to deal with at some point. This structural generalization could help explain how infants have no trouble understanding novel questions (e.g., *What did she break the stick with?*) in Study 2, which differ in surface form from questions available in their input but share the same underlying syntactic structure (19).

¹Note that preposition stranding may not always be generalized across all prepositions in a variety. For instance, while it is possible for all prepositions to undergo stranding in Lafontaine French (see Therrien, 2023), a variety of Ontario French, most other varieties of Ontario French (Roberge and Rosen, 1999) permit preposition stranding only with certain prepositions but not ones like *à* (to) and *de* (of).

(19)



(a) Structure in input

(b) Structure during test

Despite differences in linear order, toddlers successfully interpret the preposition *with* as linking back to the verb, even across an intervening verbal complement. This is consistent with findings from German-learning 19-month-olds (Hohle et al., 2006), who can also track non-local dependencies involved in *hat den Ball geholt* (*‘has the ball fetched’*) across an intervening verbal complement (*den Ball* ‘the ball’).

More broadly, infants’ readiness to understand these instrument questions before knowing the full structure may lay the foundation for identifying the more complex *wh*-dependencies shortly after in development. Prior research suggests that infants under 18 months may not yet fully represent *wh*-dependencies (Perkins and Lidz,

2021). However, our results demonstrate that even at 15 months, they can resolve various types of argument *wh*-questions. This may imply that early question understanding does not require fully mature syntactic representations. Rather, infants may respond pragmatically to partial structures of questions. In this light, their knowledge of non-local dependencies emerges shortly after their successful question comprehension at 15 months. This may be a result of their continuous effort in finding the relation between structure and meaning, along with their developing working memory (Santelmann and Jusczyk, 1998). That is, they may bootstrap their knowledge of non-local dependencies through accumulated experience with tracking and interpreting simpler structures—local dependencies, such as the links between verbs or prepositions and their neighboring arguments (Gagliardi et al., 2016; Perkins and Lidz, 2020)—and when their acoustic memory, their workspace, is robust enough to capture the non-local dependency, the complete adult-like structural representation would emerge instantly.

Alternatively, it is possible that infants already represent early forms of *wh*-dependencies at 15 months. Existing research suggests that by 12 months, infants already show sensitivity to word-order properties that differentiate declaratives from interrogatives (Geffen and Mintz, 2015). In terms of question comprehension, our work adds to prior research (Seidl et al., 2003; Gagliardi et al., 2016; Perkins and Lidz, 2020) and highlights infants’ striking capacity at 15 months for resolving questions beyond their input. This prompts us to speculate that infants at this age may have already perceived various structural signals of questions, including the fronted *wh*-phrase. Their failure to distinguish grammatical and ungrammatical *wh*-questions through listening time (cf. Perkins and Lidz, 2021) may reflect performance limitations—associated for instance with infants’ high responsiveness to questions—rather than lack of structural knowledge. Specifically, they may recognize the question

feature signaled by a clause-initial *wh*-phrase (e.g., *which dog*), and their listening time to these questions may be driven by their response to that question feature shared by both grammatical (*Which dog should the cat hug?*) and ungrammatical (**Which dog should the cat hug him?*) versions of the questions. If their attention is drawn to the question feature associated with the clause-initial *wh*-element, and their allocation of attention at this age is lopsided, this may hinder their sensitivity to the non-local syntactic dependencies—that a *wh*-phrase and a local direct object cannot co-occur in a direct object *wh*-question.

Although our findings support the transitivity-driven hypothesis, they are also compatible with an alternative explanation—15-month-olds may have succeeded by looking for unmentioned event participants, without necessarily noticing the final preposition *with*. However, this “naïve strategy” seems unlikely, considering prior research on infants’ question comprehension. First, this naïve strategy cannot explain findings from Seidl et al. (2003), where 15-month-olds succeed in responding to subject-questions (e.g., *What hit the flower?*) as well as where-questions (e.g., *Where is the flower?*). That is, infants at this age looked to the unmentioned participant only for subject-questions but not for where-questions. Second, this naïve strategy also predicts infants at this age to fail at which-questions like *Which dog did the cat bump?*, when a white dog bumps a cat and the cat bumps a brown dog. Following the naïve strategy, 15-month-olds should not show preferential looking for either the white dog or the brown dog upon hearing the question, since neither dog is explicitly mentioned in the question. This turns out not to be the case. Gagliardi et al. (2016) found that 15-month-olds look correctly at the dog intended by the question, either with a subject which-question (*Which dog bumped the cat?*) or with an object which-question (*Which dog did the cat bump*). In sum, these findings suggest that 15-month-olds’ success is unlikely to be driven by a naïve search for unmentioned

participants and instead reflects genuine sensitivity to the structural nuances of the questions.

Still, it is important to definitively rule out this alternative explanation of our results in future research. One way to achieve this is to use ungrammatical test sentences like **What did she break the stick?*, where the final preposition *with* is removed from the full grammatical question. The naïve strategy predicts that 15-month-olds would still search for the unmentioned instrument (i.e., hammer) in this case. Yet, if their correct response in Study 2 hinges on detecting a missing argument after the preposition *with*, their proportion of looking to the instrument should decline significantly with the ungrammatical version, namely in the absence of that transitivity cue—a stranded preposition *with*. To further address whether the fronted *wh*-phrase is represented for question comprehension at this age, we can use other minimally different ungrammatical test sentences like **Did she break the stick with?* In a similar vein, we would predict 15-month-olds to show the same response observed in Study 2 if they do not yet integrate the fronted *wh*-phrase during question comprehension.

Together, our findings offer initial evidence that 15-month-olds can resolve instrumental *with*-questions, despite their sparsity in infant-directed speech. This challenges the possibility that early comprehension is tightly bound to frequency or verb-specific expectations. Instead, it points to a general capacity for recovering phrase structure in early language acquisition. In interpreting questions, infants at 15 months appear to rely on cues such as argument gaps that are informed by selectional requirements of either verbs or prepositions, even when the input for these gaps is sparse. These results contribute to a growing body of work suggesting that learners before 18 months still process syntactic elements in a bottom-up fashion (Lidz et al., 2017), which supports the soon-after emergence of more abstract grammatical knowledge around 18 months (Perkins and Lidz, 2021). Future research can build on this work by explor-

ing how infants refine their representations of *wh*-dependencies, and by identifying the conditions under which predictive or top-down processing begin to shape early comprehension.

Chapter 3: Preschoolers' *the*-expressions in natural production

3.1 Background

Adult English speakers use singular definite descriptions like *the*-phrases in a way that reflects their understanding of conversational dynamics. For instance, when they use *the mug* in the sentence *Give me the mug*, their choice implies the existence of a unique mug (Russell, 1905) within a contextually restricted domain that makes the mug familiar or identifiable in the discourse or speech situation (Gundel et al., 1993; Heim, 1982; C. Roberts, 2003; Strawson, 1950). For children learning to produce these singular *the*-definites, this means acquiring the ability to assess how identifiable an intended referent is within the shared knowledge and tailoring their language accordingly. The emergence of this ability can be reflected in whether they know when to use definite forms in production.

Studies on English-learning children's acquisition of *the*-definites seem to show a puzzling asymmetry between production and comprehension. On the one hand, children aged 3 to 5 are claimed to overuse *the* in contexts expecting indefinites during elicited production, where the intended referent is either unfamiliar or non-unique (Maratsos, 1976; Schaeffer and Matthewson, 2005; Schafer and de Villiers, 2000; van Hout et al., 2010; Warden, 1976; Wexler, 2011). On the other hand, they demonstrate adult-like comprehension of *the*-definites (Aravind, 2018; Syrett et al., 2010), even before they turn age two (Choi et al., 2018; Saylor and Ganea, 2007). For

example, children aged 3 to 5 understand that definite descriptions are inappropriate when the intended referent is not unique (Syrett et al., 2010). Furthermore, toddlers as young as 19 months exhibit looking responses that presumably indicate an early grasp of the distinction between *the* and *a*—they are aware of the speaker’s visual perspective and how it might differ from their own when interpreting singular definites like *the ball* (Choi et al., 2018). These findings imply that children have the correct meanings for *the* and are sensitive to the listener’s knowledge—they seem to have the ingredients for correctly producing definite expressions.

However, there are potential issues with prior elicited production tasks. While the previous literature suggests that children systematically overuse definites in production, children’s rates of misuse vary widely depending on the experimental setup (van Hout et al., 2010). Additionally, many production studies lack adult control groups for comparison, making it difficult to define the term *overuse* precisely. Without a clear baseline, it is unclear whether children are truly overusing *the*. Moreover, the misuse of definites in elicited production may result from experimental artifacts. In fact, some studies that include adult control groups show that even adults who presumably have target knowledge of definites produce them in certain contexts where definites might be considered inappropriate (Maratsos, 1976; Schafer and de Villiers, 2000; van Hout et al., 2010).

Assuming that children do overuse definite descriptions, there are two possible explanations. On the one hand, their performance in production may be a direct reflection of their linguistic knowledge, indicating that they have not acquired the correct meaning for *the*-phrases. Alternatively, it could be that their production reflects not their linguistic knowledge, but their ability to put that knowledge to appropriate use. This hypothesis would imply either that something about previous tasks biased children away from correct performance, or that other nonlinguistic

factors might intervene between linguistic knowledge and its use in production.

In this work, we set out to test whether young children do systematically overuse *the*-definites in production, as previous literature suggests, and whether their production errors indeed reflect deficient knowledge. We examine children’s use of singular definites in natural production contexts, drawing on two corpus studies (Study 3 and Study 4), and one human simulation study with adults (Study 5).

3.1.1 The meaning of *the*-definites

Generally, speakers use *the N* when they presuppose the existence of a relevant domain of *N*s, inferrable by the audience, and refer to all its members. Thus *the mugs*, where *N* is plural, is used to refer to all the mugs in the contextually inferrable domain. When *N* is singular, as in *the mug*, the domain is taken to have a single member, and the speaker then refers to that unique *N*.

Strawson (1950) provides a useful perspective. He remarks that *the N* is used either anaphorically, when an *N* was mentioned in prior discourse, and the speaker wishes to signal reference to that same *N*, or when the context allows the listener to identify the referent, even without a prior introduction. One common theory is that such contexts are ones that exhibit *uniqueness*: only a single thing satisfies the noun *N* (Russell, 1905). This view implies that (20), for example, is acceptable only if the context of use provides a domain in which there is exactly one ball.

(20) The ball is missing.

But this theory is debatable. It can be correct only with a specialized notion of “context”. Consider (21).

(21) Mick grabbed a pebble from his collection. Then he put the pebble in a jar.

A collection of pebbles includes more than one pebble. Nevertheless, *the pebble* can be used appropriately in the second sentence of the discourse, to refer to the pebble Mick was said to have grabbed. So the “context” that would satisfy uniqueness is not the situation that contains a collection of pebbles, but rather the state of discourse created by the prior sentence. Sentences like (22) refine the required notion of “context” even further.

(22) At the gold medal match, the Russian voted for the Russian. (Neale, 2004)

Here the “context” must change within a single clause. For the subject it must be the judge, and for the object it must be the contestant. This level of subtlety has led to some to say that uniqueness is not analytically relevant. What matters is just that the intended satisfiers of N be familiar or identifiable in the discourse. We refer to all of these, which are many when N is plural, and one when it is singular (Heim, 1982; C. Roberts, 2003). Moreover, the understood domain for the description may be given only by way of defeasible inferences based on specific practical knowledge. For example, (23) is acceptable only if we assume that John’s murder was plausibly committed with a knife.

(23) John was murdered yesterday. The knife lay nearby. (C. Roberts, 2003)

What is clear, in view of this debate, is that acquiring the correct understanding of when to use *the* may not be a trivial task. Learners must infer the conditions for using *the* from rather subtle data, tracking a speaker’s referential intentions through discourse. This requires sensitivity to how the plausible domains for reference may

be shaped both by previous utterances (i.e., linguistic contexts), and by pertinent practical knowledge (i.e., pragmatic contexts) – perhaps a challenging task for young children.

3.1.2 The overuse literature

Decades of production studies suggest that this challenge of acquiring *the*-meaning is reflected in children’s errors. They found that children overuse *the*-definites up to age 5 (Maratsos, 1976; Schaeffer and Matthewson, 2005; Schafer and de Villiers, 2000; van Hout et al., 2010). These studies generally identify two broad types of errors: children misuse singular definite *the*-phrases either when the referent is unfamiliar to the listener, or when the referent is non-unique in context. Table 3.1 summarizes findings from representative studies for both error types (studies using similar methods are grouped together and separated by horizontal dashed lines).

Study	2 y.o.	3 y.o.	4 y.o.	5 y.o.	adults
Unfamiliar referent					
Warden (1976) - Exp3		54%		38%	0%
Emslie & Stevenson (1981) - Exp2	10%	8%	14%		18%
Schafer & de Villiers (2000) - Cond3		0%	0%	0%	3%
Schaeffer & Matthewson (2005)	25%				2%
Non-unique referent					
Maratsos (1974)		17%	58%		
Schafer & de Villiers (2000) - Cond8		49%	53%/41%	56%	30%
van Hout et al. (2010) - Exp1			50%		17%
van Hout et al. (2010) - Exp2			26%		24%

Table 3.1: Misuse of definite NPs for referent introduction in two- to five-year-olds

A closer look at these studies reveals the first critical pattern: children’s *the*-overuse for unfamiliar referent varies drastically depending on whether visual context

is provided. Errors for unfamiliar referents are more prevalent in tasks that incorporate visual stimuli. For instance, in one classic design, children read a cartoon story and then described it to a listener who could not see the story due to a screen in between (Warden, 1976; Emslie and Stevenson, 1981). As the listener lacked visual access to the referents, children had to introduce new entities. In another design, children observed a scene in a picture and then answered questions from a puppet who was unaware of the scene (Schaeffer and Matthewson, 2005). In this case, children had to mention referents that were old to themselves but new to the puppet (24).

(24) Situation: picture of Minnie Mouse who just finished drawing a car

Elmo: Hey, who is this?

Child: Minnie Mouse!

Elmo: And what did Minnie Mouse just do?

Child: She paint **the** car.

(Schaeffer and Matthewson, 2005: AS, 2;8)

In both cases, children frequently used *the*-definites inappropriately to introduce hearer-new referents. However, when visual support was removed, *the*-overuse disappeared entirely. Schafer and de Villiers (2000) found that children made no overuse error when referring to something hanging on the wall of their own room at home, without any visual prompt.

In contrast, errors with non-unique referents tend to emerge in tasks without visual support. In several studies, children heard a short story and were then asked to answer questions about it (Maratsos, 1974; Schafer and de Villiers, 2000; van Hout et al., 2010). For example, in Schafer and de Villiers (2000), children were asked about one of the referents in the story (25).

(25) Three ducks and two dogs were walking across a bridge. One of the animals fell off the bridge and said *Quack*. What was it?

Child: **The** duck.

(Schafer and de Villiers, 2000)

As the intended referent was non-unique, using *the*-definites would be inappropriate. Interestingly, this type of error was mitigated when visual support was provided. Van Hout, Harrigan, and de Villiers (2010) found that children made fewer errors in a story completion task when pictures were available (experiment 2) than when they relied only on verbal descriptions (experiment 1, replicating Schafer and de Villiers, 2000).

These findings raise important questions about whether previous studies have captured children’s deficient knowledge or whether experimental artifacts have inflated error rates. First, the familiarity-violating type of errors suggests that children prioritize visual context over discourse context when introducing referents (De Cat, 2015). This pattern may arise from unnatural task designs that fail to establish a natural need for perspective-taking or do not provide clear motivations for considering the listener’s knowledge. Second, the uniqueness-violating type of errors suggests that children may struggle to track referent uniqueness in discourse without visual cues. This could reflect their limitations in working memory—children may fail to maintain a coherent representation of the referential domain, leading to follow-up errors in uniqueness evaluation. Given these potential confounds, we need to critically reassess *the*-overuse errors and their implications for children’s knowledge of *the*-definites.

3.1.3 Reconsidering *the*-overuse

A careful examination of previous elicited production studies gives us reasons to question whether they truly point to a non-adult-like grasp of definites. Indeed, when we compare children’s rates of misusing *the* with those of adult control groups across different studies (Table 3.1), two key findings emerge. First, children do not exhibit consistent rates of misuse. Second, the baseline established by adult control groups is highly variable, and hence may indicate the unreliability of previous measures. Additionally, there is significant variability in misuse rates among both children and adults. To illustrate this, we plot the misuse data in Figure 3.1, which clearly shows the wide variance in both groups.¹

The wide variability in performance across both children and adults suggests that different task demands might significantly influence their referential choice. Children’s inflated rates of misuse may be tied to specific task demands, including unnatural setups for perspective-taking and unclear domains of reference.

Focusing exclusively on production studies with preschoolers may obscure the broader picture of how children’s understanding of definite descriptions evolves over time. Recent works using comprehension tasks suggest that children have an adult-like understanding of *the* from early on. For example, Syrett et al. (2010) found that when presented with two identical red balls, 3- to 5-year-old children consistently rejected or questioned an inappropriate request like *Give me the red one*, which violated the uniqueness requirement of *the*. This indicates that children associate singular *the* with some notion of uniqueness of an object within a restricted domain. Similarly, Choi et al. (2018) demonstrated that by 19 months, children are considering the

¹Abbreviations for studies: D13 = de Cat, 2013; M74 = Maratsos, 1974; Mea01 = Matthewson et al., 2001; SD00 = Schafer and de Villiers, 2000; SM = Schaeffer and Matthewson, 2005; Vea10a/b = Experiment 1/2 in van Hout et al., 2010

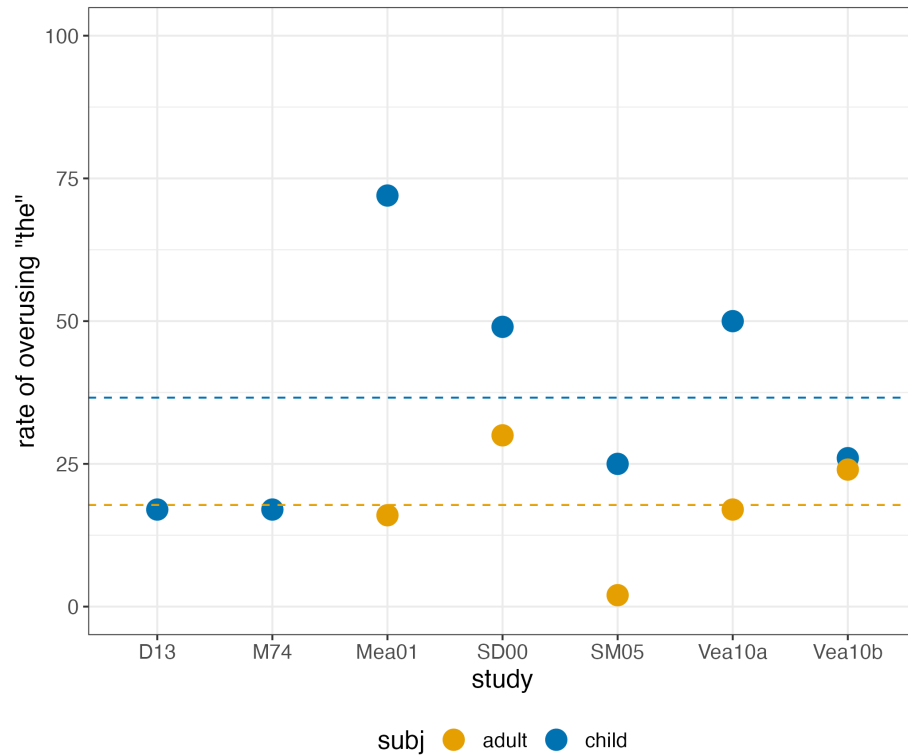


Figure 3.1: Rates of misusing the in previous elicitation studies (blue dotted line for children’s mean and orange for adults’)

speaker’s knowledge when interpreting *the*. At a scene that naturally called for visual perspective-taking, 19-month-olds, but not 14-month-olds, looked significantly longer when a listener who heard *Give me the ball* grabbed a ball hidden from the speaker, rather than one that was visible to the speaker. This was not the case when the listener heard *Give me a ball*. In fact, additional evidence indicates that the perspective-taking skills necessary for interpreting definite descriptions emerge early in life, even before children turn age two (Luo and Baillargeon, 2007; Moll et al., 2008; Saylor and Ganea, 2007).

It is important to note, however, that all evidence so far for children’s understanding of *the* is indirect. The processes involved in comprehending *the*-definites may differ from those required for producing appropriate referring expressions, so

different measures may point to different conclusions. Children might correctly respond to definites without fully considering the presuppositions associated with *the*. For instance, they might understand that a speaker can use singular *the N* to refer to some specific object—and be puzzled if they don’t understand which—without understanding that the referent should be the unique *N* in the relevant “context”. Therefore, we need more tightly controlled studies from different angles to identify which parts of the prior literature are more indicative of what children know about definite descriptions.

3.1.4 Hypotheses about children’s knowledge

The apparent production-comprehension discrepancy can be understood in two ways. Either children’s understanding of *the* is immature, and revealed in production studies, or it is adult-like, but masked in production studies.

According to the first deficient-knowledge hypothesis, children’s misuse in production reveals their immature or deficient understanding. A common explanation attributes children’s misuse to non-adult-like reasoning. Either children are egocentric, meaning they are less skilled than adults in accurately assessing what their listeners know (Karmiloff-Smith, 1979; Maratsos, 1976), or they are immature in conversational reasoning (Schaeffer and Matthewson, 2005; van Hout et al., 2010). The other theory proposes that children’s errors result from a fundamental misunderstanding of the meaning of *the*. For instance, they might have failed to grasp the uniqueness requirement (Wexler, 2011). Given these deficiencies, children’s success is limited to certain comprehension tasks, and their true, flawed understanding can only be exposed in production tasks targeting specific referential scenarios.

According to the second task-performance hypothesis, children’s production errors

are due to performance issues with artificial production setups. Unnatural task setups do not provide natural motivations for perspective-taking, while unclear referential domains lead to failed evaluation of referent uniqueness. Both of these factors may have boosted children’s inappropriate uses of *the*-expressions and eclipsed their adult-like knowledge.

To tease apart these hypotheses, it is crucial to examine children’s use of definite and indefinite expressions in both natural speech and elicited production. The deficient-knowledge hypothesis predicts that children will consistently overuse “the” in natural settings, as production will expose their fundamentally immature understanding. The task-performance hypothesis predicts that production errors should decrease significantly in more natural settings where extraneous demands are removed.

Across three studies, we demonstrate that children do not systematically overuse *the* in natural speech. Through two corpus studies (Study 3 and Study 4), we show that in natural speech, children follow their mothers’ patterns of determiner use overall and across different clause types and syntactic environments, and their rate of miscommunications due to inappropriate use of *the* is low and comparable to that in adult-adult interactions. Additionally, through an experiment (Study 5) using the Human Simulation Paradigm, we further show that when naïve adults are given only limited linguistic context, they are equally accurate at guessing determiners (*the* vs. *a*) used by children as those used by their mothers in conversation, suggesting that children’s use of determiners is adult-like.

3.2 Study 3

We begin with a corpus study to compare the frequency of singular *the*-definites produced by children and their mothers in natural speech. While it is clear that children use these referring expressions intentionally, our initial goal is to set aside the specific discourse contexts in which they occur and focus purely on how frequently *the*-definites are used. This coarse analysis helps us to assess whether there are obvious signs of children misunderstanding the meaning of *the*. Specifically, we look at the distribution of singular definites of the form *the N* used by children vs. their mothers. Do children produce singular definites more frequently than their mothers? Do children and mothers produce these definites in the same linguistic environments?

3.2.1 Method

The distributional analysis of singular definites is based on several CHILDES corpora, including Brown (1973), Soderstrom et al. (2008), Suppes (1974), and Valian (1991). These corpora provide a large sample of data, with 27 children from a broad age range (1;0-3;11), consisting of a total of 912,530 words and 254,753 sentences.

In our analysis, we focus specifically on singular definites produced by children and mothers, as these require a choice between a definite or an indefinite determiner. We analyze instances of singular definites in the form of *the N_{SG}*, while excluding plural definites, since plurals either require *the* for definite referents or no determiner for indefinite referents.

To determine whether children overuse singular definites, we use mothers' input as a baseline and compare the proportion of *the N* used by the two groups. The proportion of *the N* is calculated by dividing the count of *the N_{SG}* by the combined count of *a/an N_{SG}* and *the N_{SG}*. We exclude other determiner alternatives (e.g.,

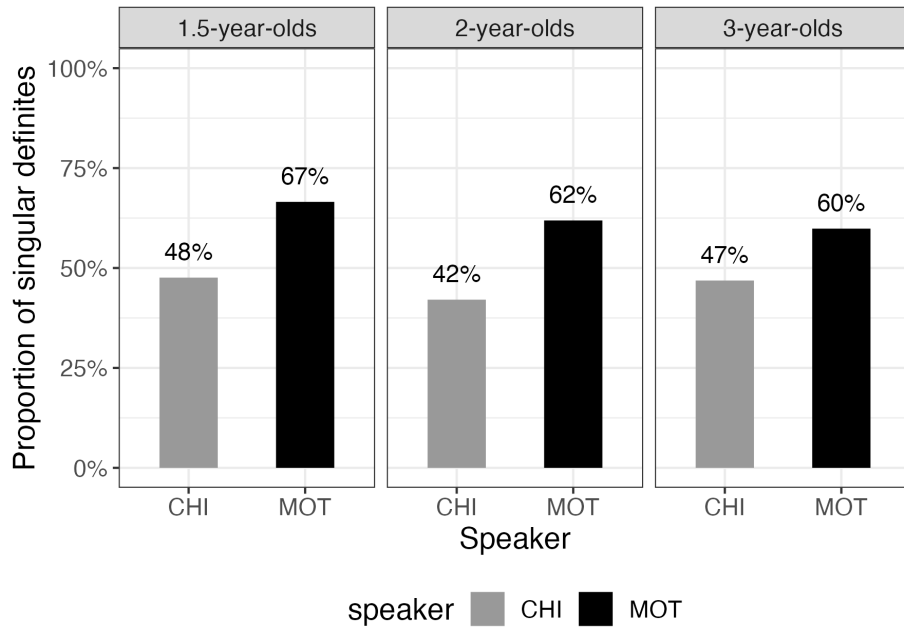


Figure 3.2: Overall proportion of singular *the*-definites by age

demonstratives like *this* and *that*) from the denominator, given that *the* and *a* are typically contrasted in elicited production studies. This analysis helps us examine whether the overuse of *the* observed in previous elicited production studies is also present in natural speech.

3.2.2 Results

In general, children use a lower proportion of singular definites than their mothers across all age groups, as shown in Figure 3.2.

Specifically, children use the definite determiner (as opposed to the indefinite determiner) about half of the time, whereas mothers use it more frequently, in over 60% of their utterances. This higher usage of *the N* by mothers may reflect their more dominant role of the relevant environments in directing the child's attention to previously mentioned or familiar objects during conversations. The absence of overuse

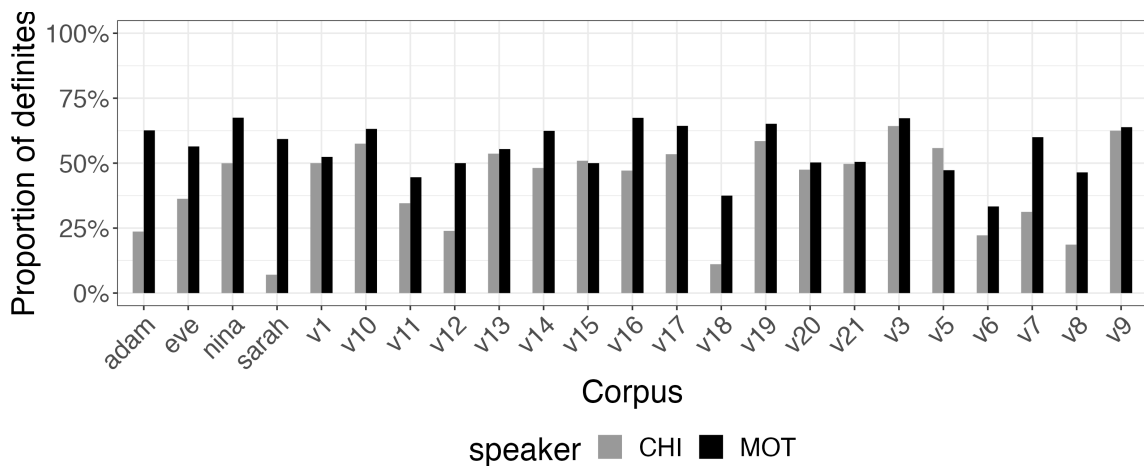


Figure 3.3: Overall proportion of singular *the*-definites by 2-year-old corpus

in children is also evident when we examine individual data from our two-year-old corpora, as depicted in Figure 3.3.

In short, this data does not provide evidence for overuse. However, the overall proportion of singular definites may be too coarse as a measure, as these differences may reflect broader differences in how adults and children participate in conversations. If this measure is not sensitive enough, what more specific measures then could we use to probe for potential overuse?

One way we can refine our measure is to examine whether children adjust their definite usage in response to pragmatic differences across clause types (declaratives vs. interrogatives). Declaratives are typically used to assert information, whereas interrogatives are used to seek information. This pragmatic difference might correlate with different probabilities for the use of definites, as well as with the different roles children and their caregivers take in conversations. So if children do overuse definites, we might expect this to be manifested differently in declaratives vs. interrogatives.

Across clause types, we did not find evidence that children overuse definites compared to their mothers' baseline usage (declaratives/statements vs. interroga-

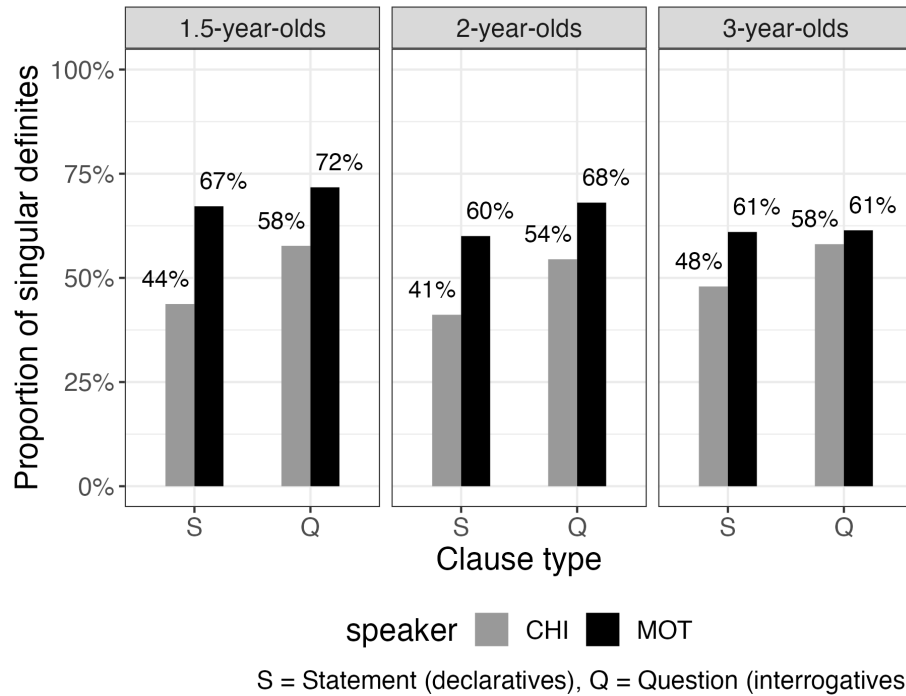


Figure 3.4: Proportion of singular the-definites across clause types (S for statements/declaratives vs. Q for questions/interrogatives) by age

tives/questions; see Figure 3.4).

In fact, children use fewer definites than their mothers, either in declaratives or in interrogatives. This means that even when asserting information, where a referent is often already known or inferable within the discourse, children do not overuse definites in referring to established entities.

Another way we can refine our measure is to examine whether children adjust their definite usage in response to pragmatic differences across syntactic environments (subjects vs. objects vs. fragments) for declaratives. In English, subjects tend to refer to entities mentioned earlier in the discourse (Prince, 1992), so we expect caregivers to use more definites in subject positions. If children were truly overusing definites, we would expect their proportion of definite usage to be insensitive to such distinctions. To test this, we conducted further analyses to examine whether this finer-grained

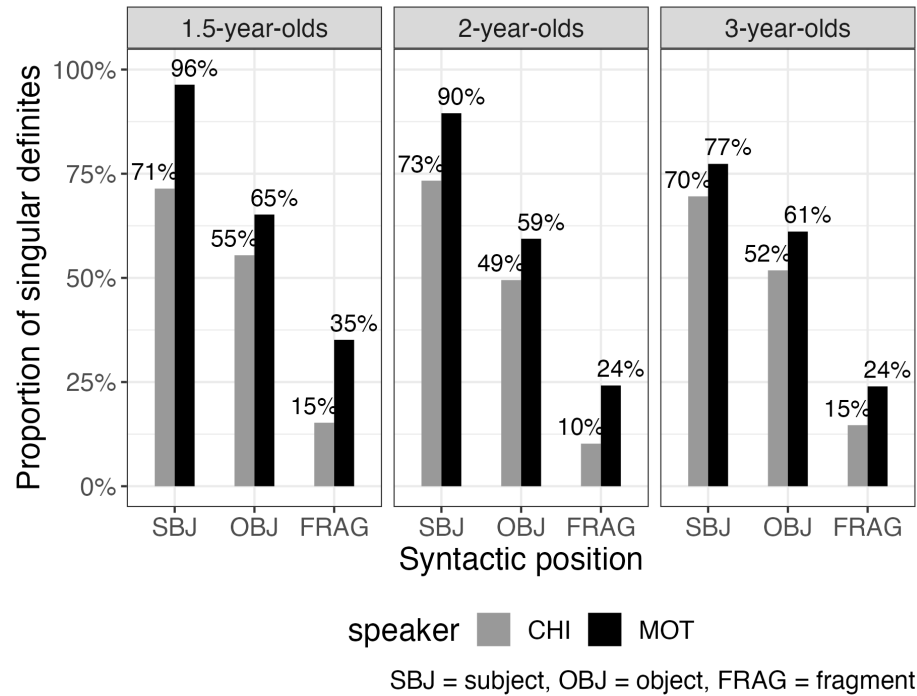


Figure 3.5: Proportion of singular *the*-definites across syntactic positions

measure would reveal children’s overuse of definites.

Across syntactic positions (i.e., subjects, objects, and fragments) in declarative clauses, we again find no evidence of overuse when examining children’s determiner use (see Figure 3.5).

Moreover, children use more singular definites for subjects than for objects, mirroring their mothers’ patterns. That is, children’s usage patterns align with the general observation that in English, subjects typically refer to entities mentioned earlier in the discourse (Prince, 1992).

One concern is that children’s rate of definites might be underestimated relative to their mothers’ due to young children frequently producing noun phrases (NPs) without determiners (e.g., Hyams, 1996). If children tend to drop determiners more often for definites than for indefinites, excluding these null determiners could significantly distort estimates of their overall rate of definite usage. To explore this possibility, we

identified instances of null determiners in children’s speech (see Table 3.2).

Age (year)	Def	Indef	Null	(Def+Null)%
1	209	166	550	82.1
2	1,364	1,342	2,599	74.7

Table 3.2: Revised rate of *the*-definites (excluding fragment NPs)

If we were to assume that all of these null or omitted determiners were meant to be definites, the revised rates of definites for 1-year-olds and 2-year-olds would rise to 82.1% and 74.7%, respectively—surpassing the rates observed in their mothers’ speech, i.e., 66.5% and 60.2% in the same corpora. However, this assumption likely overestimates children’s use of definites, as some of these null determiners are clearly used as indefinites, whether grammatically (26) or not (27).

(26) I want coffee.

(Valian, 1991; 01;09;25)

(27) I take walk.

(Brown, 1973: Eve, 01;09;00)

To obtain a more accurate estimate, we used children’s observed distribution of definites across syntactic positions to infer the likelihood that a null determiner in a particular position was intended as a definite. Specifically, we recalculated children’s revised definite rates by probabilistically including null determiners that could plausibly be definites, i.e., $N_{\text{null as def}}$ in (29). This calculation was based on the observed rates of definites produced in subject ($\text{rate}_{\text{def-sub}}$) and object ($\text{rate}_{\text{def-obj}}$) positions (28). In other words, by considering how frequently children’s definites occur in these syntactic environments, we can better estimate the proportion of omitted determiners that likely reflect intended definites.

$$(28) \quad N_{\text{null as def}} = N_{\text{null-sub}} \times \text{rate}_{\text{def-subj}} + N_{\text{null-obj}} \times \text{rate}_{\text{def-obj}}$$

$$(29) \quad \text{rate}_{\text{revised def}} = (N_{\text{def}} + N_{\text{null as def}}) / (N_{\text{def}} + N_{\text{ind}} + N_{\text{null}})$$

The revised rates of definites for 1- and 2-year-olds then become 58.2% and 52.1%, respectively, both lower than their mothers' rates of 66.5% and 60.2% in the respective corpora. Therefore, when using this finer-grained estimate, children do not appear to overuse definites in their natural production.

3.2.3 Discussion

In summary, we find no clear evidence of overuse in children's natural speech. Across different clause types and syntactic positions, children's distribution of singular definites aligns closely with that of their mothers.

While these broad frequency patterns suggest that children do not overuse definite determiners, they may mask more subtle forms of overuse. At early stages, children may be simply mimicking adult speech, without fully grasping the underlying rules—producing definites in the right places, but not for the right reasons. In other words, a lack of overuse does not necessarily imply adult-like knowledge. Importantly, overuse of definites in frequency is not an inevitable sign of limited knowledge, but one possible outcome among others. We need to move beyond surface frequencies and examine how definites are used in actual communicative contexts.

To this end, the next study shifts focus to children's use of definites in discourse contexts, miscommunications in particular. Overuse may only become apparent when considering how children's definites function in communication. By analyzing instances of miscommunication, we aim to identify whether children may be using definite descriptions in ways that do not align with the referential expectations of

their listeners—offering a more sensitive test of potential overuse.

3.3 Study 4

In this study, we evaluate the quality of children’s use of *the*-definites. If children misunderstand the meaning of *the*, we would expect them to use it inappropriately, leading to communication breakdowns. Specifically, if children use *the* in contexts where its presuppositions were not satisfied or could not be accommodated, their listeners would likely be confused about the referent or seek clarification.

To assess whether miscommunications induced by *the* are common in natural speech, we compare the rates of such miscommunications in mother-child interactions with those in adult-adult interactions. We calculate the rate of misuse of *the* by 1- and 2-year-olds and compare it to that of adults. If children’s understanding of *the* is non-adult-like, we would expect to see a higher rate of miscommunications, with children’s misuse of *the* frequently causing listener confusion or prompting requests for clarification.

3.3.1 Method

The rate of *the*-driven miscommunications was analyzed using CHILDES corpora, including Brown (1973), Soderstrom et al. (2008), Suppes (1974), and Valian (1991), while the rate for adult-adult interactions was calculated using the CALLHOME corpus (Kingsbury et al., 1997).²

For children, we focused on miscommunications involving singular definites, where

²The CALLHOME corpus consists of 120 unscripted, 30-minute telephone conversations between native English speakers, primarily involving family members or close friends. This is particularly relevant because the corpora we selected for analyzing children’s miscommunication rates also involve interactions between mothers and their children, who are familiar with each other. The CALLHOME corpus includes a total of 301,805 words and 28,967 sentences.

the choice of determiner is between *the* and *a/an*. For adults, we expanded the scope to include both singular and plural definites to capture a more comprehensive view of their miscommunication scenarios. In the analysis of mother-child interactions from CHILDES, sentences containing *a/an N_{SG}* or *the N_{SG}* were extracted using CLAN commands and functions.

For the analysis of adult-adult interactions in the CALLHOME corpus, sentences with referring expressions were extracted using a Python script.³ To identify cases of *the*-driven miscommunications, we searched for question marks, as a sign for clarification, within 1 to 3 lines following the target line containing the definite expression. We then manually reviewed these cases to confirm that they were true instances of miscommunication. To do so, we expanded the context and checked if they genuinely involved clarification questions.

3.3.2 Results

We first analyzed the use of *the N* by 1-year-olds and found only 1 instance of misuse out of 205 cases (0.49%), with no instances of miscommunication. Many of the uses of *the N* by 1-year-olds were repetitions of their mothers' utterances or followed the mothers' initial mention of *N*. After excluding these 14 repeated cases from the total, the misuse rate remained low at 0.52% (1 out of 191).

Next, we examined the use of *the* by 2-year-olds and found 0.39% (10 out of 2,575) instances of miscommunication driven by *the*. We identified two types of misuse: unfamiliar reference (8 cases) and non-unique reference (2 cases). (30) shows a case of unfamiliar reference, and (31) a case of non-unique reference (see Appendix C for the full set of misuse cases observed in 2-year-olds' corpora).

³Many thanks to Weihang Wang and Sathvik Nair for helping with the script.

- (30) CHILD: I wanna go on the platform.
MOTHER: What platform?
CHILD: The platform at the circus.
(Valian, 1991: 02;04;26)
- (31) CHILD: Hey, where's the truck? I need ...
INVESTIGATOR: Where's the what?
MOTHER: The truck?
CHILD: Yeah.
MOTHER: There's two trucks.
(Brown, 1973: Eve, 02;02;00)

In order to assess whether these cases reflect a semantic or pragmatic deficit, we further examined adult-adult miscommunications to determine whether this rate of miscommunication or the kinds of miscommunications differ between children and adults.

In adult-adult interactions, we found 0.13% (4 out of 3,121) instances where *the* led to miscommunications. All these cases were due to unfamiliar reference. (32) is one example of this type of case.

- (32) A: You know, spent panting and spending every second of the day, w- with him only. You know what I'm saying and that probably part of the reason that it would be nice to actually stay with him is that **you could also go and visit the sights** or whatever and g- and...
B: Where? What are you talking about?
A: In Japan.

3.3.3 Discussion

In summary, Study 4 provides no clear evidence that children overuse *the* in natural production. The rate of miscommunication driven by *the* in mother-child interactions is low and comparable to that in adult-adult interactions, both being well below 1%. This low rate of miscommunication contradicts the deficient-knowledge hypothesis, which predicts that children should overuse *the* in natural production.

However, the absence of miscommunication observed in Study 4 may be due to mothers accommodating their children’s use of definites. Mothers tend to know their children well, and the lack of clear signs for miscommunication may be owing to a more predictable common ground for mothers to identify the intended referents. To evaluate children’s use of determiners more objectively, we need to involve adults who were not part of the conversations and thus are less likely to accommodate the presuppositions.

3.4 Study 5

In Study 5, we aim to determine whether children’s use of determiners is adult-like by conducting a determiner-guessing experiment, using Dieuleveut et al.’s (2022) adaptation of the Human Simulation Paradigm (Gillette et al., 1999). Dieuleveut et al. show that this method is sensitive to whether children are adult-like in choosing between a necessity modal (like *must*) or a possibility modal (like *might*). We use the method here, supposing it will be likewise sensitive to the choice between a definite and an indefinite description.

We presented adult participants with excerpts from mother-child conversations and asked them to select either definite *the* or indefinite *a* for a missing determiner

used by either a child or a mother. Our first goal is to determine if adults can accurately guess the definites used by mothers. If they can correctly predict mothers' use of definites, it indicates that the context, combined with adults' understanding of determiners, provides sufficient information for accurate identification of the determiners that mothers used. Conversely, if children misuse definites, adults should perform worse at predicting children's determiners. Assuming that children's overuse is characterized by using definites in indefinite contexts, then the contexts in which children use definites would likely elicit indefinite choices by adults.

If children overuse definite determiners, we expect adults to be less accurate in predicting children's use of definites compared to predicting mothers'. On the other hand, if children do not overuse definite determiners, we expect no significant difference in adults' accuracy of guessing for children's versus mothers' use of definites.

3.4.1 Method

Participants were recruited via Amazon Mechanical Turk (AMT), and the experiment was conducted using PCIBex farm (Zehr and Schwarz, 2018). The task took approximately 15 to 20 minutes. Each participant completed 40 trials, including 10 trials with definites, 10 with indefinites, and 20 fillers. To ensure participants were engaged, one-fifth of the trials (8 trials) were followed by comprehension questions (i.e., simple memory questions). See Appendix D for the full lists of stimuli for participants in different conditions.

The materials consisted of snippets from mother-child conversations involving 23 pairs (children aged 2;1;4 to 4;11;2). These snippets were randomly selected from Gleason's corpora in CHILDES (80,347 words; Menn & Gleason, 1986).⁴ To better

⁴The conversations took place in recorded lab sessions, where children and mothers interacted with new objects. This type of interaction provides a nice setup to probe participants' use of

capture cases of misusing *the* if any, we ensured that both *the* and *a* were viable options for their subsequent noun so that the participants’ choice of either determiner would be based solely on contextual information. To do this, we excluded items biased toward either determiner. Items were removed if they included: 1) repetitions, such as echoing, speaker disfluency, and finishing another speaker’s sentence; 2) idiomatic expressions (e.g., *wait a minute*) or collocations; or 3) plural or uncountable nouns.

Before starting the experiment, participants received instructions (see Figure D.1 in Appendix D for the full instructions participants received) and two practice trials to familiarize themselves with the process. They learned that: 1) they needed to press the spacebar to reveal each line of conversation in progression, 2) they would choose between two options for the blank at the end of each conversation, and 3) they would occasionally answer comprehension questions to ensure they understood the context. Figure 3.6 shows a sample test trial, where participants chose a determiner from options including target determiners (e.g., *the* vs. *a*), filler determiners (e.g., *this* vs. *that*), or mixed determiners (e.g., *the* vs. *this*).

240 participants (127 males and 105 females; mean age: 41.3) were recruited via AMT. Data from 211 participants (109 males, 94 females, and 2 unspecified; mean age: 41.9) were included in the analysis. Participants who scored below 75% on the comprehension questions were excluded from the final dataset.

3.4.2 Results

We analyzed the rate of correct guesses for mothers’ use of definites as a baseline to evaluate children’s use of definites. Adults predicted mothers’ definites above chance for conversations from all age groups (2-year-olds: $V = .666$, $p < .001$; 3-year-olds: $V = .666$, $p < .001$).

 definites, as we are interested in exploring how children register new referents in a dynamic common ground.

CHILD: Yeah..
MOTHER: Yeah?
MOTHER: I'll go get it.
MOTHER: Here, let's put the car back where it belongs.
MOTHER: Okay.
CHILD: No!
MOTHER: You don't wanna put the car there?
MOTHER: Okay.
CHILD: No.
CHILD: Wanna play with it and read ___ book, okay?

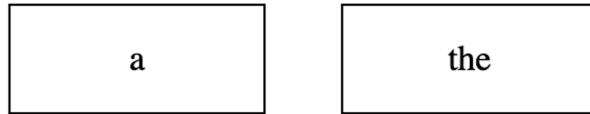


Figure 3.6: A sample test trial during Study 4

$V = 780, p < .001$; 4-year-olds: $V = 389, p < .001$).⁵ Their accuracy rates were 0.9 ($N = 36, SD = .096$) for conversations with 2-year-olds, 0.895 ($N = 39, SD = .102$) with 3-year-olds, and with 4-year-olds, 0.765 ($N = 31, SD = .236$).

Adults also predicted children's definites above chance for conversations from all age groups (2-year-olds: $V = 595, p < .001$; 3-year-olds: $V = 595, p < .001$; 4-year-olds: $V = 595, p < .001$). Their accuracy rates were 0.844 ($N = 36, SD = .144$) for 2-year-olds, 0.857 ($N = 35, SD = .127$) for 3-year-olds, and 0.894 ($N = 34, SD = .110$) for 4-year-olds. Crucially, in no age group were adults significantly better at predicting mothers' definites than children's (2-year-olds: $W = 516, p = .124$; 3-year-olds: $W = 571.5, p = .212$; 4-year-olds: $W = 726.5, p = .007^{**}$) (Figure 3.7).⁶

⁵Here, V represents the sum of ranks for one group in a Mann-Whitney U test (also known as the Wilcoxon rank-sum test) often used to compare differences between two independent groups. It is a non-parametric test, without assuming normality.

⁶The Wilcoxon test was used to test whether adults were above chance at guessing definites for

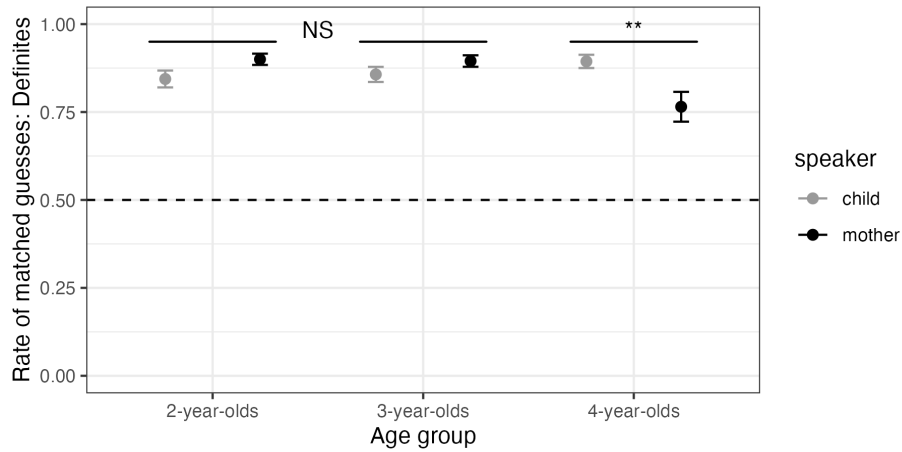


Figure 3.7: Rate of correct guesses for definites by speakers and child age

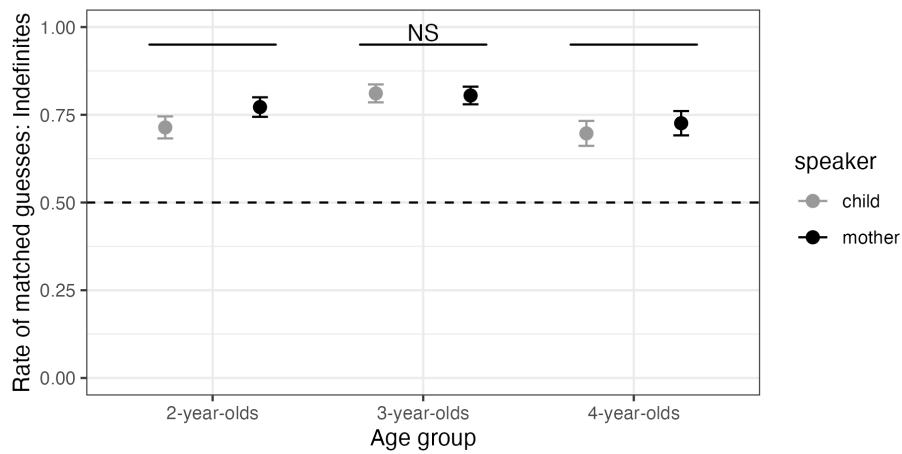


Figure 3.8: Rate of correct guesses for indefinites by speakers and child age

One might suspect that adults' accurate predictions could be due to a general bias toward choosing definites, which would suggest poorer performance with indefinites. However, this is not supported by our data (Figure 3.8).

Adults were also above chance in predicting indefinites used by both children (2-year-olds: $V = 593.5$, $p < .001$; 3-year-olds: $V = 527$, $p < .001$; 4-year-olds: $V = 511.5$, $p < .001$) and mothers (2-year-olds: $V = 587$, $p < .001$; 3-year-olds: $V = 768$,

each condition, and the Mann-Whitney U test was for testing whether adults performed better with either child or mother utterances. These non-parametric tests were used, as the data did not follow a normal distribution.

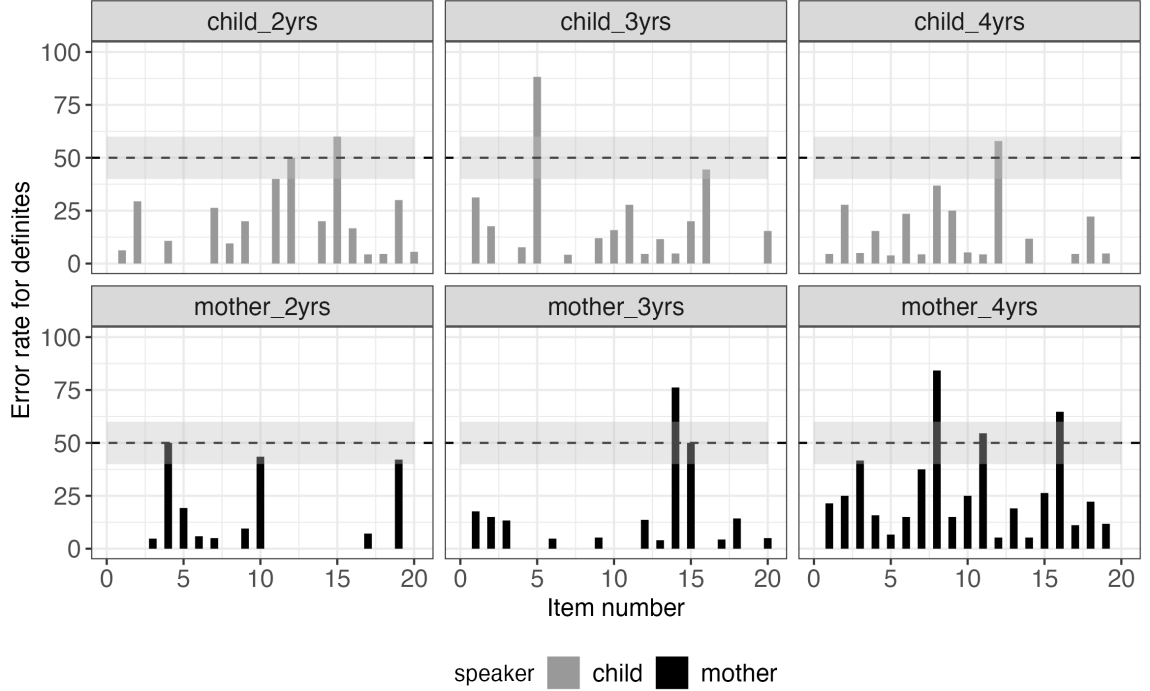


Figure 3.9: Adults' error rates for definites by items, speaker, and child age

$p < .001$; 4-year-olds: $V = 385$, $p < .001$). Furthermore, adults were not significantly better at predicting mothers' indefinites compared to children's (2-year-olds: $W = 526$, $p = .163$; 3-year-olds: $W = 703.5$, $p = .819$; 4-year-olds: $W = 498$, $p = .704$).

Additionally, participants generally showed agreement in their determiner choices for most test items, whether the determiners were used by children or mothers (Figure 3.9). We used gray to mark regions with an accuracy range between 40% and 60% to highlight items with significant disagreement (i.e., a near 50-50 split among participants). The figure shows very few items with such wide disagreement. Our examination of items with the highest mismatch rates revealed only one potential case (item 5 for child determiner uses in 3-year-old corpora) of misusing *the* by children (see Appendix D for item indices and the corresponding test materials).

In summary, our results suggest that children use singular definites in a man-

ner consistent with adult usage in natural production, as evidenced by adults' high accuracy in identifying definites used by both children and their mothers.

3.5 General discussion

Across three studies, we observed that in natural production, children often use referring expressions in ways similar to adults. In natural speech, children generally do not use singular definites more frequently than their mothers; instead, their use of definites and indefinites aligns closely with their mothers' across various clause types and syntactic environments. Notably, the rate of miscommunication driven by *the*-definites in child-mother interactions is extremely low, comparable to that observed in adult-adult conversations. Additionally, naïve adults are equally accurate in guessing *the* or *a* used by children and their mothers, given snippets of mother-child conversations.

Our findings lend preliminary support to the task-performance hypothesis. The overuse of *the*-definites found in elicitation studies may be due to experimental artifacts. If children's understanding of *the*-definites is deficient, we should be able to observe signs of overuse. However, we found that children's uses of definites in natural conversations are adult-like, with only occasional usage errors in their computing of the referential domain. The absence of systematic errors in natural production suggests that children's knowledge of definites is in place but may be liable to interference from artificial contexts.

We propose two general explanations for children's tendency to overuse the definite determiner *the* in some elicitation tasks. For their overuse errors for unfamiliar referents, one possible explanation is that children rely more on visual context than on discourse context when introducing referents (Karmiloff-Smith, 1979; De Cat,

2015). This tendency likely stems from task designs that do not naturally encourage perspective-taking or fail to clearly prompt children to consider the listener’s knowledge. In terms of their overuse errors for non-unique referents, it could be that children may have difficulty assessing referent uniqueness without visual support. This challenge may be linked to their working memory constraints. Without potential referents in sight, children might struggle more to maintain their representation of the referential domain solely with linguistic description, compared to adults (van Hout et al., 2010). This degraded maintenance of the referential context in memory for evaluating referent uniqueness may be responsible for children’s increased errors observed in prior studies (Schafer and de Villiers, 2000).

While natural production provides valuable insights into children’s language use, it has certain limitations. During mother-child interactions, the common ground tends to be less dynamic and more predictable than during an experimental setting, which might make it easier for children to assess what their listener knows. In other words, children’s adult-like usages might be limited to mostly easy cases in natural production, so natural production scenarios might be insufficient to expose underlyingly deficient knowledge.

To sum up, we show through three studies that 1) there is a lack of clear evidence for systematic *the*-overuse in children’s natural production, and 2) their use of *the* is adult-like. Children’s determiner use follows their mothers’ overall and across clause types and syntactic environments; their rate of *the*-driven miscommunications is low; adults are equally successful at guessing definites used by children and their mothers. Our results suggest that children follow the discourse well, and *the*-overuse reported in elicitation studies may reflect performance issues due to experimental artifacts.

Chapter 4: Overcoming performance issues in elicited production

4.1 Background

Chapter 3 focused on children's use of *the*-definites in natural conversations. I provided new evidence that children use *the*-definites in an adult-like way in daily scenarios, where the artificiality of experimental settings is removed.

However, natural conversations with caregivers may be insufficient for us to rule out the possibility that children lack full knowledge of *the*-definites. First, the common ground between children and caregivers tends to be relatively stable, with fewer updates in shared knowledge compared to experimental settings. Because caregivers often share knowledge with their children, for children to assess what caregivers know might be easier than assessing what others know. Second, daily scenarios may not be variant enough to expose underlyingly deficient knowledge. For instance, if most referents in natural interactions are uniquely identifiable, children may not need to actively evaluate the uniqueness of referents when choosing their referential forms.

For children to produce appropriate referring expressions, they must (i) understand the presuppositions of different referential forms, (ii) consider the perspective of their interlocutor, and (iii) assess, in real time, whether the intended referent is uniquely identifiable within a contextually defined referential domain. If any of these components are missing or disrupted, production errors may arise.

Prior production studies indicate that children overuse *the*-definites, but the pre-

cise cause for these errors remains unclear. As discussed in chapter 3, unnatural setups and unclear referential domains may have contributed to children’s overuse of *the*-definites in prior studies. The goal of this chapter is to investigate whether children have acquired adult-like knowledge of *the*-definites, when the other two components of perspective-taking and referential domain are controlled for in a task.

By speculating about the task-performance hypothesis, one important premise we entertain is that children have the capacity for both perspective-taking and identifying referential domains. To determine what type of task we need in probing children’s referential abilities in production, we must first review what we currently know about their cognitive capacities in early development.

4.1.1 Development of perspective taking

Perspective-taking is a fundamental cognitive skill that underlies successful communication. It allows individuals to consider others’ viewpoints, knowledge states, and attentional focus, which is essential for producing effective referring expressions. From infancy through the preschool years, children gradually refine their ability to interpret and respond to others’ perspectives, laying the groundwork for more sophisticated social interactions, such as the production of referring expressions.

Early signs of perspective-taking emerge in infancy. Research suggests that by 12 months, infants can track others’ visual perspectives. For example, Ikeda, Kanakogi, and Hirai (2022) found that 12-month-olds could follow an adult’s gaze accurately when the line of sight was clear, whereas 7-month-old did not. Furthermore, infants at this age not only recognize what others see but also use this information to interpret others’ actions (Luo and Baillargeon, 2007) and provide information through gestures based on an adult’s knowledge state (Liszkowski et al., 2008).

As children transition into toddlerhood, they begin to integrate perspective-taking into their understanding of language. By 12 to 18 months, infants track others' prior experiences when interpreting their requests. Tomasello and Haberl (2003) showed that 12- and 18-month-olds could consider others' experience in identifying intended referents. In their study, infants played with two adults and two new toys, but one of the adults left the room while the child and the other adult played with a third toy. When the adult returned and saw all three toys aligned on a tray, she requested *Wow! Cool! Look at that one! Can you give it to me?*, and infants successfully retrieved the toy new to the speaker. Similarly, Saylor and Ganea (2007) found that 14- to 20-month-olds considered others' prior experiences in resolving ambiguous requests. In their task, two experimenters played separately with a different ball before putting it in a particular box. When one experimenter later made an ambiguous request *Where's the ball?*, infants correctly chose the intended ball based on prior interaction with the speaker, against a baseline condition without such request.

By 19 months, toddlers exhibit more sophisticated visual perspective-taking in understanding communication. Choi, Song, and Luo (2018) demonstrated through a looking-time task that 19-month-old, but not 14-month-old infants expected an addressee who heard a request like *Give me the ball* to reach for a ball that was visible to, not hidden from, the speaker. However, it is not until around 24 months that toddlers consistently integrate visual perspective-taking into their behavior. Moll and Tomasello (2006) found that 24-month-olds, but not 18-month-olds, successfully handed the experimenter a toy occluded from his view (but not theirs) upon hearing *Where is the other toy*.

By age two, children begin to overcome egocentric biases. Lempers, Flavell, and Flavell (1977) found that 2-year-olds could orient pictures in an adult fashion for another person to see—they understand that others have different visual perspectives.

This awareness extends beyond the visual domain (Moll et al., 2014), and by age three, children exhibit Level 2 perspective-taking, understanding that objects can look different to others depending on their viewpoint (Moll and Meltzoff, 2011).

Another crucial aspect of perspective-taking is children’s ability to attribute mental states (e.g., beliefs and desires) to others, known as Theory of Mind (or ToM). While classic false-belief tasks suggest that ToM emerges around four years of age (Hogrefe et al., 1986; Wimmer and Perner, 1983), more recent studies show that younger children can succeed in modified tasks. 3-year-olds perform better on false-belief tasks when pragmatic demands are reduced (Freeman et al., 1991; S. Lewis et al., 2017), and even 15-month-olds succeed in attributing false belief in a nonverbal task (Onishi and Baillargeon, 2005).

In addition to pragmatic demands, children’s deployment of perspective-taking skills is influenced by task demands on executive function (EF). Children’s ability to suppress their own perspective correlates with their EF skills (Carlson and Moses, 2001; C. Hughes and Ensor, 2007) and their application in communication (Nilsen and Graham, 2009). For instance, Nilsen and Graham (2009) found that in comprehension, 3- to 4-year-olds’ ability to inhibit their own perspective was directly related to their inhibitory control. More specifically, their performance on conflict EF, to flexibly inhibit and activate different responses, has been found to relate uniformly to ToM measures, independent of verbal ability (Carlson et al., 2015). Moreover, the link between EF and ToM seems to be specific to ToM tasks that require children’s reasoning about others’ epistemic states (Henning et al., 2011).

In sum, children’s perspective-taking abilities emerge early and continues to be refined and adapted for its use in different settings. Task demand is a factor that directly influences children’s deployment of perspective-taking. While nonverbal perspective-taking is evinced in infants and toddlers, performance in verbal tasks varies signif-

icantly depending on the specific settings. In the next section, I spell out other cognitive demands that may impact children’s referential encoding.

4.1.2 Other demands on encoding reference

Most production studies investigating the use of *the*-definites focus on children over age two (Schaeffer and Matthewson, 2005; Maratsos, 1974; Schafer and de Villiers, 2000; van Hout et al., 2010), meaning that their perspective-taking abilities should already be in place. Yet, children still face challenges in referential production, suggesting that additional cognitive demands may be at play.

First, children may not spontaneously engage in perspective-taking unless motivated. In non-verbal tasks, infants and toddlers take others’ perspectives to predict eye gaze (Ikeda et al., 2022), interpret interactions (Choi et al., 2018), and engage socially (Moll and Tomasello, 2006). However, production tasks, such as storytelling, lack direct interactions, which may reduce children’s motivation to take another’s perspective. Even adults are more likely to engage in perspective-taking when given explicit incentives (Epley, Keysar, et al., 2004). This suggests that perspective-taking is not an automated process but one that is contextually motivated.

Second, cognitively costly tasks can impair children’s accuracy in selecting appropriate expressions. Studies show that adults, too, exhibit egocentric tendencies in production (Fukumura and van Gompel, 2012) and comprehension (Epley, Morewedge, and Keysar, 2004), particularly under cognitive load (Vogels et al., 2013). When inhibiting their own perspective for referent identification is costly, adults rely more on their own perspective (Brown-Schmidt, 2009). Similarly, children’s misuse of *the*-definites may stem not from egocentrism, but from cognitive overload.

In light of this, instead of attributing children’s errors simply to egocentricity, we

should identify task contexts where errors arise. One such context is picture-based storytelling, where children tend to misuse *the*-definites for unfamiliar referents (Emalie and Stevenson, 1981; Warden, 1976). Beyond encoding reference, these tasks impose additional cognitive demands: children must systematically relate pictorial representations to their referents (see Ganea et al., 2009, DeLoache and Burns, 1994), interpret picture sequences as temporally coherent events (see Cohn, 2020), and, last but not least, consider common ground by distancing themselves from the immediate visual context (see Carlson et al., 2015). These extra layers of complexity may contribute to children’s production errors in narrative tasks.

A related challenge emerges in story completion tasks with unclear referential domains, where children tend to use *the*-definites for non-unique referents (Maratsos, 1974; Schafer and de Villiers, 2000; van Hout et al., 2010). Early studies of this kind (Schafer and de Villiers, 2000; Maratsos, 1974) lacked visual support, requiring children to track both the types and tokens of verbally introduced referents. This can be demanding on children’s working memory. Notably, even adult control groups produced *the*-definites in 20% to 30% of cases in such contexts (see Condition 4 and 8 in Schafer and de Villiers, 2000). Later studies (van Hout et al., 2010) introduced picture-based stimuli, which aided referent tracking, but as these images were visually dense, and the referential domain varied in space, it can be harder for children to establish a clear referential domain. Indeed, misuse persists in adult control groups for about a quarter of the cases (see Experiment 2 in van Hout et al., 2010).

However, there is evidence that in experimental settings, children can incorporate perspective-taking into referential communication as early as age three. At this stage, they show some sensitivity to their interlocutor’s perspective in speech production, with noticeable refinement by age four (Matthews et al., 2006; Salomo et al., 2011). For instance, Matthews et al. (2006) found that 3- to 4-year-olds adjusted their use

of referring expressions, i.e., full nouns vs. pronouns, based on the givenness of the intended referent reflected in its perceptual availability and prior discourse. When their interlocutor could see the intended referent or had recently mentioned it, children favored pronouns over full nouns; when these referent-given cues were absent, they showed the opposite pattern by favoring full nouns over pronouns.

Taken together, these findings highlight that children’s errors with *the*-definites cannot be reduced to a failure in perspective-taking alone. Instead, they arise in contexts that introduce additional processing demands—whether in mapping visual information to linguistic forms, maintaining referents across a verbally presented narrative, or resolving referential domain in cluttered visual scenes. Understanding how these factors interact is crucial for disentangling children’s linguistic competence from task-specific performance limitations.

4.2 Study 6

The task-performance hypothesis predicts that children’s misuse of *the*-definites will decrease significantly in a production task where task demands introduced by unnatural setups and unclear referential domains are controlled for. To test this prediction, we conducted Study 6, in which we adopt a conversational elicited production paradigm, aiming to create a more natural setup and a clear domain of reference.

4.2.1 Method

Participants in this task need to consider the perspective of the listener and assess whether the intended referent is uniquely identifiable within the domain of mutually known referents. The task setup involves Kermit the frog, who cannot see objects on the other side of a wall, while the participants can, so Kermit needs some help from

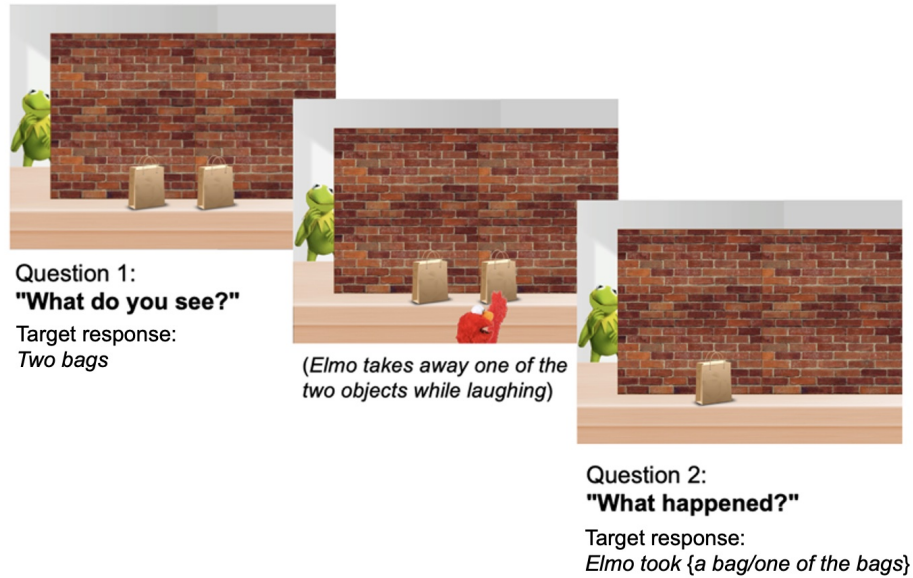


Figure 4.1: Experimental setup for Study 6

participants to tell him what's behind some walls (Figure 4.1; see Appendix E for the details of the complete setup and procedure). Sometimes, two identical objects are behind the wall, which should prompt participants to use an indefinite to refer to one of the objects. Sometimes, two distinct objects are behind the wall, which should lead participants to use a definite to refer to one of the objects.

Before the task starts, the experimenter and the participant sit in front of a laptop. The experimenter explains that Kermit the frog needs some help from the participant and will be very happy if the participant can help him. This serves to motivate the participant to be a cooperative interlocutor. Throughout the task, the experimenter uses a presentation remote clicker to control the slides on the screen, thereby reinforcing the impression that the conversation is spontaneous and natural.

To ensure that participants understand the procedure and are comfortable with the setup, we begin with three practice trials before the main task, each involving just one object behind the wall. In the first two trials, Kermit only asks Question

1, *What do you see*. After the participant responds, Kermit says, *Okay, thanks*. The third trial differs from the previous two in that after the participant answers Question 1, Elmo appears, takes away the object, and laughs. Kermit then asks, *Wait. Did you hear something? It sounded like someone's behind the wall. Who was it*. This question ensures participants understand the situation. Finally, Kermit asks Question 2, *What happened*, so that the participant knows what question to expect during the test trials. Once the practice trials are over, the experimenter explains that he has some work to do and will not join the conversation, and then he turns sideways to prevent the participant from being distracted. This is to ensure that during production, the participant will solely engage with Kermit and consider Kermit's perspective.

The test trials include two trial types, where participants see two objects, either unique (e.g., a bottle and a mug) or non-unique (e.g., two identical bags). Kermit first asks, *What do you see* (Question 1). We expect this to elicit indefinite responses (e.g., *a bottle and a mug* or *two bags*) regardless of the trial type (i.e., unique vs. non-unique trial), since the referents are unfamiliar to the listener (Kermit) and not within the shared visual common ground between Kermit and the participant. Next, Elmo appears on the other side of the wall and takes away one of the objects, while laughing. Kermit then asks, *What happened* (Question 2). Here, we expect participants to respond with definites for unique referents (e.g., *Elmo took the mug* if there's only one mug) but indefinites for non-unique referents (e.g., *Elmo took a bag/one of the bags* if there are two bags). However, if participants are egocentric (Maratsos, 1976) or mistakenly assume shared knowledge (Schaeffer and Matthewson, 2005), they would ignore that the listener cannot see which object has been taken away for the non-unique referents (e.g., two identical bags). Consequently, they are expected to incorrectly use a definite description (e.g., *the bag*) for a particular referent that they

themselves know, which would replicate the pattern of misusing *the*-expressions for non-unique referents observed in the literature (Schafer and de Villiers, 2000; Maratsos, 1976). Two orders of test trials are created, with the unique and the non-unique test trials pseudo-randomized.

Our design reduces task demands in the following aspects. First of all, our task involves turn-taking in elicitation, which mimicks natural conversations and helps maintain the natural flow of speech. Additionally, as Kermit cannot see what’s behind the wall and there are no referents in the visual common ground, it’s clear to participants that the task requires visual perspective-taking, which has been proven to work even with infants (Brezack et al., 2021; Choi et al., 2018; Luo and Bailargeon, 2007). Moreover, we create a clear domain of reference (the area behind the wall) and provide visual support for the referents, so the memory load is minimized. Furthermore, we control for the saliency of the intended referent. When Question 2 (*What happened*) is asked, Elmo has already taken the intended referent away, so it’s not visually salient, and discourse-wise, *What happened* also serves as a neutral prompt from an ignorant interlocutor (Kermit), targeting the broader context (i.e., what occurred behind the wall) rather than the specific identity of the stolen object.

We employ a 2 (Trial type: unique vs. non-unique) $\times$ 2 (Question: 1 vs. 2) $\times$ 2 (Age: adult vs. children) within-subjects design. If children’s knowledge of *the* is adult-like, we expect their use of definites to vary based on Trial type and Question, without any interaction effect with Age. Responses coded as definites include pronoun *it*, *the N* (e.g., *the box*), *the Adj N* (e.g., *the left box*), and *the N PP* (e.g., *the box on the left*), and responses coded as indefinites include *a/an N* (e.g., *a box*), *one N* (e.g., *one box*), and *one of the Ns* (e.g., *one of the boxes*). See Table E.4 in Appendix E for the full list of twelve referential contexts.

The participants involved 34 children (15 males) between the ages of 3;4;15 and

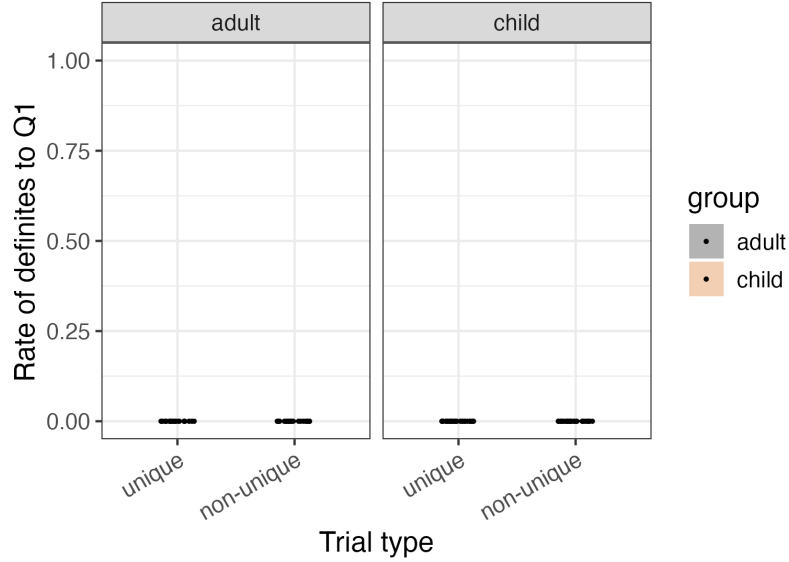


Figure 4.2: Rate of definites for Question 1 (adults vs. children)

5;0;23 (mean: 4;4;17). They were recruited from preschools in the greater Washington, DC, area and were included in the analyses only if they heard English during at least 80% of their waking hours. We analyzed data from trials where they responded to both questions (*What do you see* and *What happened*). We recruited another 23 undergraduate students at University of Maryland, College Park as adult controls.

4.2.2 Results

Our results indicate that children ($N = 34$), just like adults ($N = 23$), respect the presuppositions of *the*. For Question 1 (*What do you see*), both adults and children always used indefinites in their responses, when the intended referents were unfamiliar to the listener (Figure 4.2). For Question 2 (*What happened*), they mostly used definites (89.2%, $SD = .262$; adult control: 99.3%, $SD = .035$) to refer to a unique referent and much less so (28.3%, $SD = .368$; adult control: 27.5%, $SD = .416$) for a non-unique one (Figure 4.3).

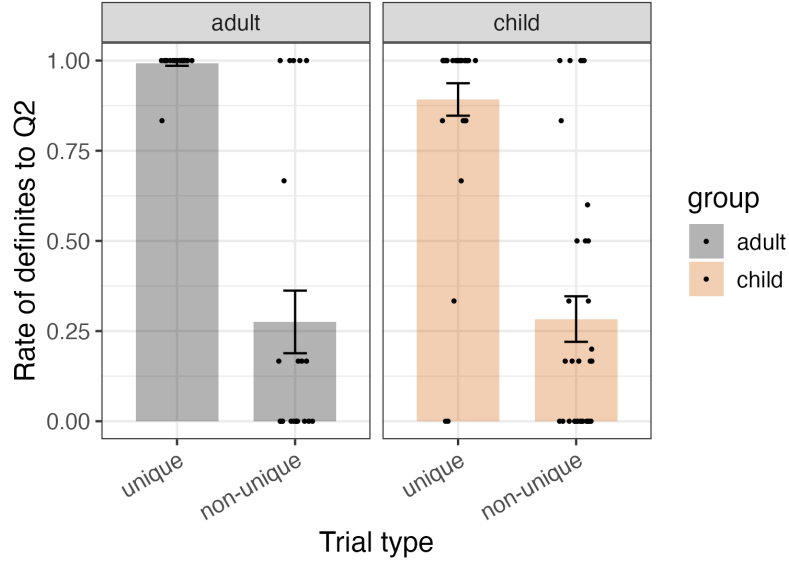


Figure 4.3: Rate of definites for Question 2 (adults vs. children)

A $2 \times 2 \times 2$ mixed ANOVA confirms that children as a group behave as adults in their rate of definite responses. We find significant main effects of Trial Type ($F(1, 55) = 59.8, p < .001$) and Question ($F(1, 55) = 154.365, p < .001$) but not Age. The interaction is significant between Trial Type and Question ($F(1, 55) = 59.8, p < .001$) but not with Age. To check whether there is an age difference among children, we test whether children's rates of definite responses correlate with their age in days. Our results suggest no such correlation for either unique referents (Figure 4.4), $r = 0.19, p = .27$, or non-unique referents, $r = -0.24, p = .168$ (Figure 4.5).

Children's specific responses to Question 1 and 2 confirm that their choice of referring expressions depends on both the type of the question and the uniqueness of the referent. For Question 1 (*What do you see*) asking children to introduce two unfamiliar referents, they predominantly used singular forms for unique referents but plural forms for non-unique ones (Figure 4.6). For Question 2 (*What happened*) asking children to refer to one of the two referents, children used *the*-definites 82.2% of the time (162 out of 197 utterances) for unique referents. However, this usage dropped

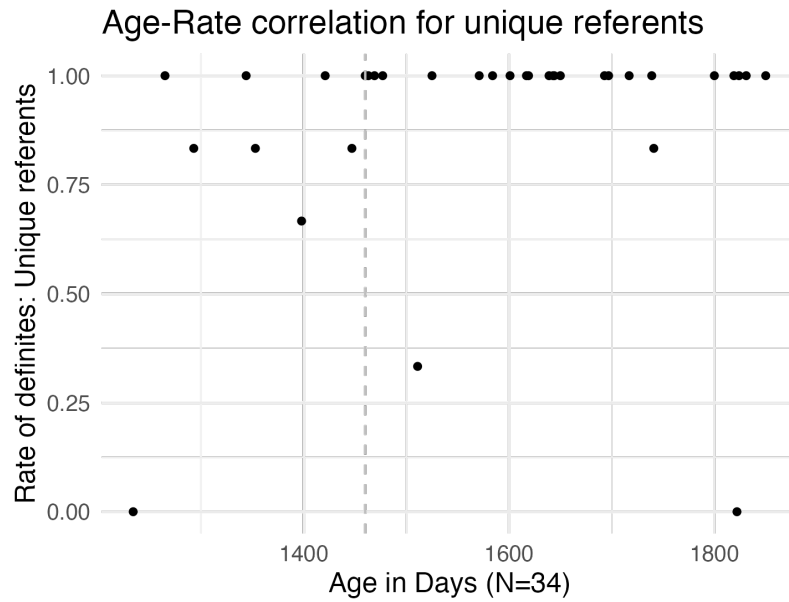


Figure 4.4: Rate of definites for unique referents by age in days (dashed line separates the data points for 3-year-olds and 4-year-olds to indicate group distinction)

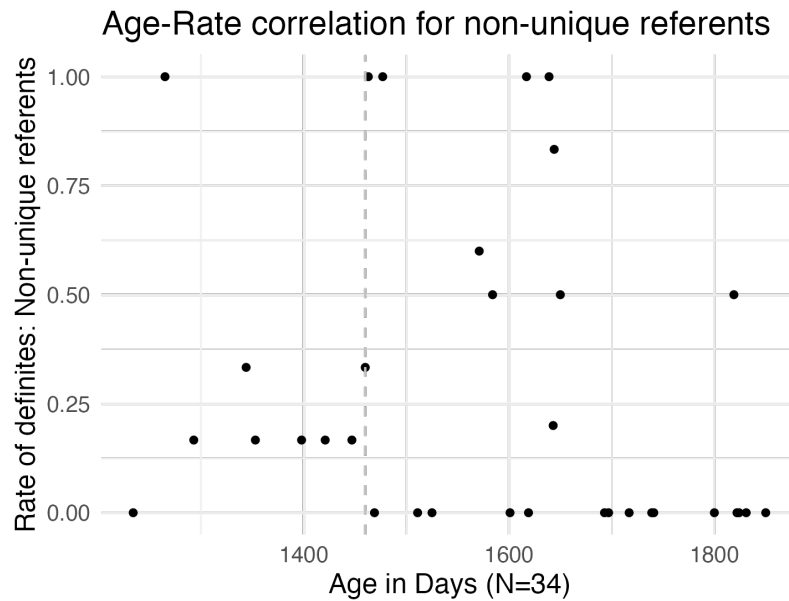


Figure 4.5: Rate of definites for non-unique referents by age in days (dashed line separates the data points for 3-year-olds and 4-year-olds to indicate group distinction)

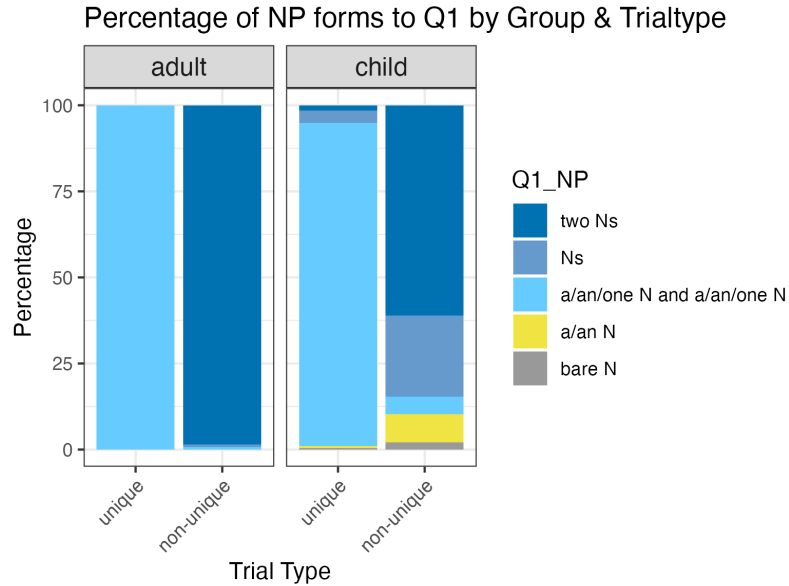


Figure 4.6: Percentage of NP forms for Question 1 by Group and Trial Type

to 13.8% (27 out of 195 utterances) for non-unique ones (Figure 4.7).

When we look closer at individual data, systematic misuse of *the N* was limited only to 3 children (out of 34) in the sample (Figure 4.8).¹ This suggests that most children know that using *the N* is inappropriate to pick out non-unique items.

In sum, our results suggest that like adults, children respect the presuppositions of *the*-expressions and use *the*-definites appropriately in an elicited production task with better controlled setup and domain of reference than in previous studies. They consistently use indefinites to introduce two unfamiliar referents to a conversation, and when referring to one of the two referents, they use *the*-definites only when the intended referent is uniquely identifiable within the domain of reference.

¹Unlike other children who produced much fewer errors, these three children showing systematic overuse of *the*-expressions seemed to represent narrower domains of reference already in the initial setup, as indicated by their responses to Question 1, in scenarios with two non-unique objects (e.g., *two boxes*), in which they used either a singular NP (e.g., *a box*) or conjoined NPs (e.g., *a box and a box*). This is consistent with the idea that failure in identifying the target domain maybe responsible for children's overuse of *the*-expressions in prior studies.

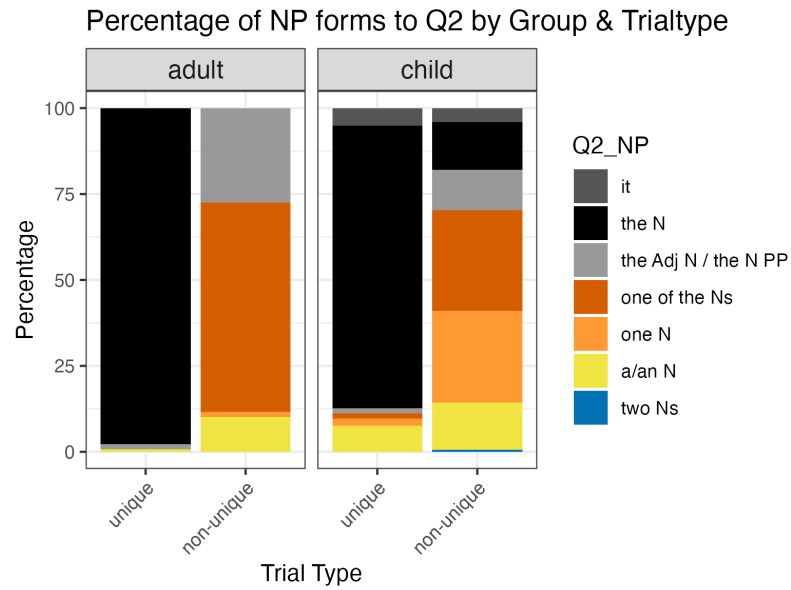


Figure 4.7: Percentage of NP forms for Question 2 by Group and Trial Type

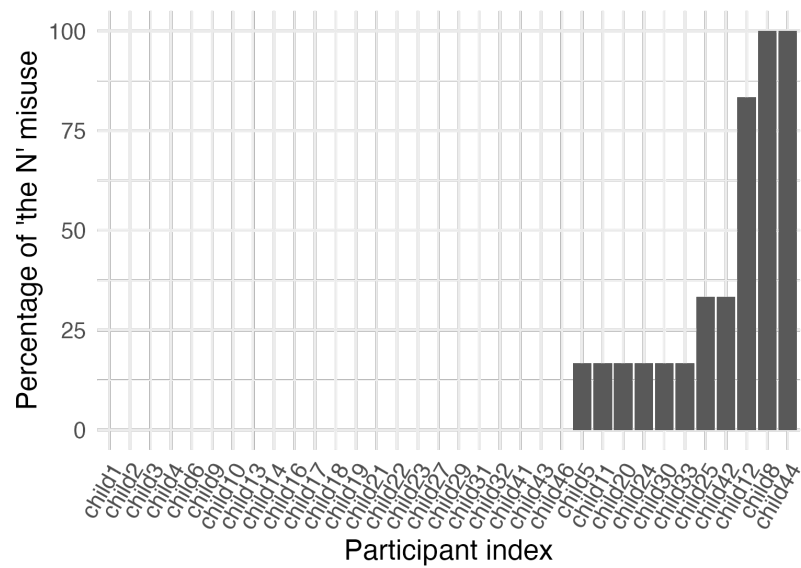


Figure 4.8: Percentage of misuse of *the N* by individuals for non-unique referents

4.3 General discussion

Our findings across four studies offer converging evidence that children consistently use *the*-definites in an adult-like manner. Despite our best efforts, we found no systematic misuse of singular definites in children’s production in either natural or elicited settings, contrary to previous claims about overuse (Maratsos, 1976; Schaeffer and Matthewson, 2005; Schafer and de Villiers, 2000; van Hout et al., 2010; Warden, 1976; Wexler, 2011). Crucially, in an elicited production task with a better controlled setup and domain of reference, even 3- and 4-year-olds behave adult-like in producing referring expressions, using *the*-definites only when the intended referent was uniquely identifiable within the domain of reference.

Our findings support the task-performance hypothesis and caution against over-interpreting production errors observed in children as evidence of flawed knowledge. When comparing the rates of misusing *the*-expressions in our study to those in the literature, we found that our rates of misuse were the lowest both for unfamiliar referents (Figure 4.9) and for non-unique referents (Figure 4.10).

Previous research has often attributed children’s misuse of singular definites across studies to systematic immaturity, suggesting that children are egocentric (Karmiloff-Smith, 1979; Maratsos, 1976), linguistically immature (Matthewson et al., 2001; Schafer and de Villiers, 2000; van Hout et al., 2010; Wexler, 2011), or pragmatically incompetent to differentiate between interlocutors’ knowledge states or to integrate shared knowledge into their use of referring expressions (Schaeffer and Matthewson, 2005). However, what has been largely overlooked is the influence of task demands (De Cat, 2015; Gagarina and Bohnacker, 2022), which might have caused children to behave differently than they would in more natural setups and even led to variable performance in adult control groups.

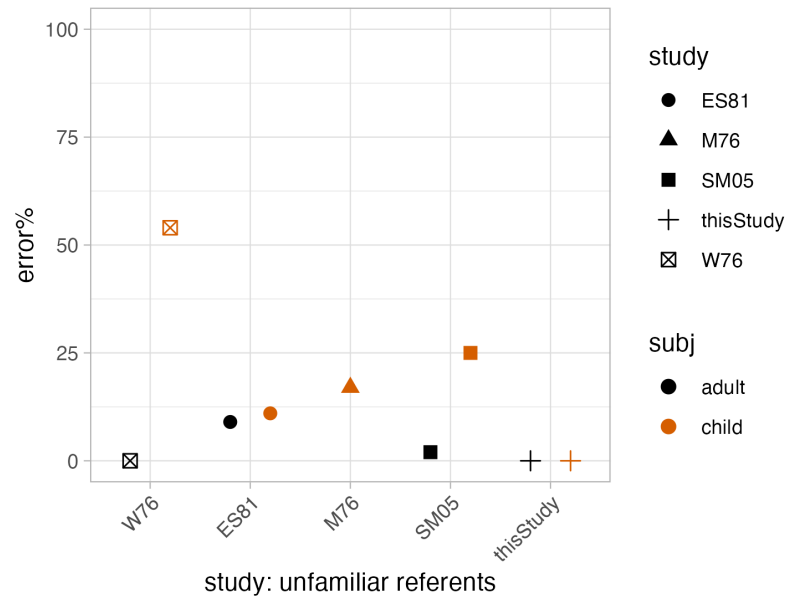


Figure 4.9: Percentage of *the N* errors for unfamiliar referents across studies

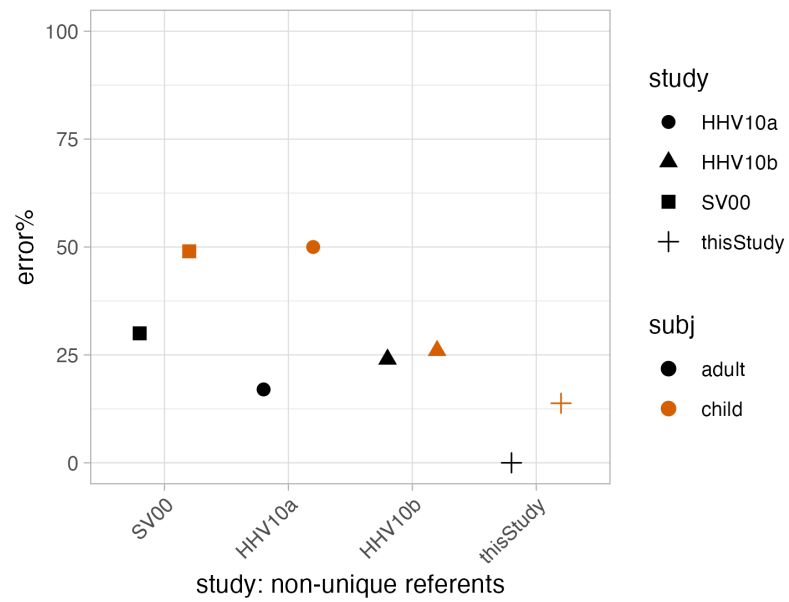


Figure 4.10: Percentage of *the N* errors for non-unique referents across studies

Our findings highlight the importance of natural setups and clear referential domains in assessing children’s competence with definite descriptions. Tasks such as story telling (Emslie and Stevenson, 1981; Warden, 1976) and story completion (Maratsos, 1974) may inflate children’s error rates for unfamiliar referents, as they are less interactive and do not necessitate perspective-taking. In contrast, our task is designed to encourage perspective-taking by establishing a clear and consistent speaker-hearer dynamic, with visual perspective as an integral component. With these factors controlled for, we find that even 3-year-olds successfully consider the listener’s perspective when selecting determiners. Similarly, tasks like question-after-story without visual support (Schafer and de Villiers, 2000) or story completion with pictures (van Hout et al., 2010) may have contributed to children’s errors with non-unique referents. These errors could stem from a disrupted referential domain—children may have difficulties maintaining types and tokens of referents in discourse due to limitations in verbal memory (van Rij, 2012) or correcting an attentional state skewed towards irrelevant visual elements (Trueswell et al., 2011). To reduce these demands, our task provides a clear and reliable referential domain (i.e., the area in front of a wall), ensuring that children focus on relevant referents. Moreover, by restricting visual access to just two objects (e.g., two identical boxes) within that domain, we minimize memory demands and prevent distractions for children. With these concerns addressed, our results suggest that like adults, 3- to 4-year-olds consider the unique identifiability of the intended referent when producing definite descriptions.

When considered more broadly, our findings underscore an important lesson: while children have the necessary pragmatic competence, their ability to demonstrate it can be hindered by the demands of a given task. When key elements of genuine communication—such as a clear addressee or a motivation to be a cooperative speaker—are absent, children might be less inclined to fine-tune their referring

expressions to the listener’s needs, as doing so requires additional effort. However, Study 5 shows that their performance improves significantly in natural communicative contexts. This pattern is not unique to definite descriptions. In comprehension, 2-year-olds demonstrate adult-like understanding of pronouns when engaged in a conversational context (Moyer et al., 2015). In production, children could adjust the informativeness of their event descriptions based on what their listener could see, but only when both the child and the listener were actively engaged in a collaborative game (Grigoroglou and Papafragou, 2019). Similarly, improving the pragmatics of a task has been shown to bring out preschoolers’ competence with complex structures like relative clauses in both comprehension (Hamburger and Crain, 1982) and production (Crain et al., 1990). Furthermore, this boost in performance with pragmatically sensible contexts also extends beyond language to cognitive abilities such as Theory of Mind (Buttelmann et al., 2009) and altruism (Warneken and Tomasello, 2006; Warneken, 2016) during infancy. Together, these findings highlight the critical role of task design in assessing children’s pragmatic competence, emphasizing the need for interactive and socially meaningful contexts to reveal their full knowledge.

Our research also provides new insights into the development of the functional vocabulary in children, particularly in relation to their cognitive abilities. Children’s understanding of determiners develops quickly, even though this requires some understanding of conversational dynamics. By 11 months, they have perceived the forms of determiners (Shi et al., 2006) and are sensitive to their co-occurrence with nouns (Babineau and Christophe, 2022); a few months later, they use them to categorize novel nouns (Mintz, 2006); by 19 months, they have learned to use speaker knowledge to restrict the domain of reference (Choi et al., 2018; for earlier evidence with 15-month-old multilinguals, see Liberman et al., 2017). Our work adds another piece to the puzzle, showing that by age two, children demonstrate adult-like us-

age of *the* and *a* in natural production, and by age three, they show adult-like use of referring expressions in elicited production. They are capable of considering the knowledge state of their listeners and adapting their referential choices accordingly across various contexts. Unlike previous studies that report frequent misuse of singular *the*-definites by children, we observed a very low incidence of such misuse in our production study. Even when examining individual data, these errors were confined to very few participants. This suggests that the pragmatic competence required to navigate presuppositions triggered by *the*-definites, often subtle and backgrounded, is present from an early age for most children, consistent with what was found in other domains of language, such as belief verbs (Dudley, 2017) and belief reports (S. Lewis et al., 2017). This also aligns with other research indicating children’s early ability to engage in perspective-taking, both in non-linguistic (Luo and Baillargeon, 2007) and linguistic contexts (Saylor and Ganea, 2007).

Taken together, our work provides consistent evidence that children produce singular *the*-definites in an adult-like manner, both in natural and elicited speech. To do so successfully, they must assess the knowledge states of their listeners and select an appropriate referential form for listeners to identify the intended referent within a specific context. By creating a natural setup and a clear domain of reference, we find that even 3- to 4-year-old children demonstrate an adult-like understanding of *the*-definites. This implies, contrary to prior overuse claims in the literature, that there is little reason to believe that children have the wrong meanings for *the*-definites or lack the pragmatic capacity to use it properly.

Chapter 5: How experience matters in learning *the*-expressions

5.1 Background

Learning when and how to use *the*-expressions is far from straightforward for young language learners. The definite determiner *the* signals that a referent is uniquely identifiable and mutually familiar to both speaker and listener—an abstract meaning that is not directly observable in the input. For children to acquire this meaning, they must identify signals, linguistic or non-linguistic, in their early environment that reliably correlate with definiteness. This chapter explores the nature of those cues in early input and considers what kinds of inferential processes children need to engage in to arrive at an early understanding of *the*. By examining the signals available to toddlers, we aim to shed light on how children begin to master one of the more subtle aspects of reference in language—how the meaning of the determiner *the* helps narrow down and identify the speaker’s intended referent within a given context.

The definite determiner *the* can be used broadly in two ways: anaphorically, referring back to a co-referential antecedent ($X_i \dots the \textit{dog}_i$), or instead situationally, identifying referents presumed to be otherwise “familiar” within the shared context (Strawson, 1950). Anaphoric uses might provide surface signals in the distribution to *the*-meaning, e.g., through repetition of nouns (*a ball...the ball*) (Fisher and Tokura, 1995; Sneed, 2007). Learners need to represent discourse referents and infer the fa-

miliarity presupposition by observing anaphoric uses of *the*-expressions for discourse referents already established through prior mention. But situational uses depend wholly on understanding the discourse, context, and speaker intent. Learners must track referential contexts or situations and infer the presupposition through observing hearer-old uses of *the*-expressions for familiar referents licensed by contextual or world knowledge. Here, inferential processes based on these two different uses are summarized in Table 5.1.

inferential process	cue	referent
anaphoricity-driven	<i>the</i> -expressions used anaphorically for repeated nouns in subsequent mentions	discourse-old
situation-driven	hearer-old use of <i>the</i> -expressions in first mentions (when the speaker expects the listener to identify the referent given the situation)	contextually given

Table 5.1: Inferring the presupposition of *the*-expressions through different uses

This chapter addresses whether the surface patterns of noun repetition found in anaphoric uses are sufficient for young children to learn the meaning of *the*. To do this, we evaluate two hypotheses: the anaphoricity-driven hypothesis and the situation-driven hypothesis.

The anaphoricity-driven hypothesis suggests that learners can infer *the*-meaning from anaphoric uses by simply attending to easily noticeable distributional features of discourse segments. As prior research (Sneed, 2007; Fisher and Tokura, 1995) has shown, definite descriptions are more likely than indefinite ones to occur when a referent is mentioned for the second time. Conversely, indefinite descriptions tend to decrease significantly in second mentions as opposed to first mentions. If children possess an innate sensitivity to the categorial distinction between definites and indefinites,

this surface difference in distribution may serve as a clear, easily observable signal for learners to bootstrap *the*-meaning through noun repetition—that *the*-expressions are anaphoric and thus encode definiteness. In this view, children don’t need to reason about the speaker’s intent and situational dynamics to understand the implicit familiarity underlying reference; rather, they can learn that *the* typically picks out referents already mentioned—something old in the discourse—just by noticing how *the* often appears with nouns they have already heard.

The situation-driven hypothesis posits that toddlers must use situational understanding to link *the*-expressions to referents that are mutually familiar upon first mention. Under this hypothesis, inference about speakers’ intention underlying reference serves as a crucial path for learners to identify the meaning of *the*. They can use situational cues—such as physical presence, shared visual attention, or goal-directed behavior—to infer whether a referent is likely to be mutually known. Even at around 12 months, toddlers begin to exhibit an emerging capacity for pragmatic reasoning in social interactions (Tomasello and Haberl, 2003; Saylor and Ganea, 2007). This raises the possibility that learners at this age may be sensitive to situations in which *the*-expressions are used. They might notice that speakers use definite descriptions for referents that are uniquely identifiable or taken for granted in a given situation—even when those referents are new to the discourse, i.e., previously unmentioned. Crucially, in inferring what *the* means, learners need to contrast linguistic alternatives, i.e., *the*-expressions vs. *a*-expressions, where the latter, indefinites, typically introduce new or non-specific referents. This contrastive reasoning allows them to pragmatically enrich their interpretation of definite descriptions, even in the absence of prior linguistic mention. Thus, the situation-driven hypothesis frames the inferential process as being necessarily grounded in both social cognition and sensitivity to linguistic alternatives.

To evaluate these hypotheses about how early learning of *the*-meaning can be achieved, we need to first examine the early input infants actually receive. This approach allows us to ask what kinds of information or cues are available and usable across scenarios, and whether they are sufficient for inferring target meanings.

For availability, distributional cues for distinguishing between *the* and *a* may not always be present across contexts. We need to consider whether learners might arrive at consistent generalizations, given signals from first vs. subsequent mentions across varying real life scenarios. As prior studies focused on specific interactive scenarios such as picture description (Sneed, 2007) and staged event description (Fisher and Tokura, 1995), it is yet unclear whether distributional cues signaled by anaphoricity are sufficient for learners to generate the target meaning across different scenarios of caregiver-child interaction.

For usability, infants' intake—the subset of linguistic data they actually process and retain (Lidz and Gagliardi, 2015; Omaki and Lidz, 2015)—may not be the same as the input. As everyday conversations unfold rapidly and often incur competing demands on attention, young learners may miss crucial elements of discourse, such as the first mention or introduction of a referent. What looks as a second-mention to a corpus annotator may, from the child's perspective, be perceived as a first-mention, especially if their attention has shifted in the interim. This disconnect means that traditional corpus analyses based only on first- and second-mentions may overestimate the portion of evidence usable by learners. Young learners are subject to their capacity of working memory (Saylor and Ganea, 2007) and attentional span (Ruff and Lawson, 1990) in reliably tracking discourse referents. To truly understand the viability of anaphoricity-based inference of *the*-meaning, we need independent evidence for young learners' robust capacity for tracking as well as maintaining discourse referents.

Additionally, anaphoric uses of *the*-expressions in the input may compete with

other uses that do not rely on prior mention in the discourse. For instance, even in adult-directed speech, a substantial portion of *the*-phrases are used for referents that are new (around 50%) to the discourse (Poesio and Vieira, 1997)—namely ones that have not been previously introduced—as opposed to ones used for old referents (around 30%). This suggests that anaphoricity, while potentially useful, may not be the signature property learners attend to when forming hypotheses about the meaning of *the*-expressions.

If learners focus simply on anaphoric or distributional cues, frequent use of *the*-expressions for first-mention referents would be puzzling. Specifically, the frequent use of *the* for discourse-new referents cannot be easily reconcilable with the generalization that *the*-expressions are more commonly used for discourse-old referents. Such input would be unclear for learners, as it is compatible with another non-target possible meaning—*the*-expressions are used for intended referents that are specific to the speaker. This speaker-specific meaning is possible for definite expressions in St’át’imcets (Schaeffer and Matthewson, 2005) for instance, where the same article (*ti* ... *a*) is used for a referent believed by both the speaker and the hearer (33) as well as one believed by the speaker only (34).¹

- (33) ka hál’h-a **ti** snéqwem-a
 OOC show-OOC DET sun-DET
 ‘The sun appeared.’

- (34) ka hál’h-a **ti** nkaúksent-a
 OOC show-OOC DET star-DET
 ‘A star appeared.’

Crucially, this kind of non-target speaker-specific meaning may arise in early development, when learners try to rationalize *the*-expressions used for first mention referents

¹OOO = out of control

but in the wrong direction, ignoring their requirement on the listener’s knowledge. If so, we would observe some learners infelicitously using *the*-expressions for first-mention referents, unaware of the requirement on mutual familiarity.

As a result, we are left with several open questions. Are anaphoric uses of *the* in early input sufficient for learners to infer *the*-meaning? What other uses of *the* coexist or even compete with anaphoric uses in early input? How do learners navigate this variability? Do different learners end up acquiring different interpretations based on mixed uses of *the*, or do they converge on a unified meaning for *the*, despite the variability in the conditioning environments of *the*-expressions?

This chapter draws on longitudinal video corpora of 12 caregiver-child pairs (Snow et al., 1996). We focus on toddlers’ input from caregivers at 14 months (Study 7) and their own output at 20 months (Study 8). Our findings reveal that anaphoric cues to the meaning of *the*-expressions are not clear-cut, whereas situational cues are consistently available in the input. Moreover, we show that by 20 months, toddlers’ spontaneous use of *the* reflects situational understanding. They do not assign *the*-expressions an inappropriate speaker-specific interpretation—that is, they are not using *the* for entities known only to the speaker but not shared with the listener (contra non-target speaker-specific interpretations possible in other languages, see Schaeffer and Matthewson, 2005; Rozendaal and Baker, 2008). Instead, they use *the*-expressions in contextually appropriate ways, demonstrating sensitivity to both discourse and broader contextual knowledge.

By considering the availability, usability and consistency of different cues to *the*-meaning in early input, this chapter aims to offer a more nuanced account of how learners may acquire the familiarity presupposition of *the*-expressions. We show that even under variable conditions, young learners are equipped to extract meaningful generalizations about the familiarity presupposition of *the*-expressions, as evinced

by their earliest use of *the* around 20 months. This highlights the importance of situational reasoning for learners to recognize the presuppositions of *the*-expressions despite the indeterminacy of distributional cues.

5.1.1 Sensitivity to discourse vs. contextual knowledge

A speaker's choice of nominal forms for a referent depends on its givenness in the discourse and contextual knowledge between interlocutors. A definite description (*the ball*) typically signals that a referent is mutually familiar, either through discourse/anaphoricity, i.e., being previously mentioned, or through perceptual availability given the situation. In contrast, an indefinite description (*a ball*) can be used for unfamiliar referents. This distinction highlights how speakers need to attend to the identifiability of the intended referent in choosing referring expressions.

Several studies suggest that in production, children show earlier sensitivity to discourse-given referents than to contextually given referents (Campbell et al., 2000; Matthews et al., 2006; Rozendaal and Baker, 2008). In natural speech, Rozendaal and Baker (2008) found that 2- to 3-year-olds frequently use definite and demonstrative determiners as opposed to indefinites for discourse-given referents, yet there is little evidence that they use definites for contextually given referents. Similarly, in elicited speech, Matthews et al. (2006) observed that while 2-year-olds do not yet use lexical nouns vs. pronouns differentially based on the perceptual availability of the referent in the experimental context, they do exhibit some sensitivity to prior discourse, but not significantly so as 3- and 4-year-olds (for similar findings, see Campbell et al., 2000).

On the other hand, evidence from natural speech suggests that children are already sensitive to contextual knowledge in using referential forms at the earliest stage.

For instance, Salazar Orvig et al. (2010) found that French-speaking children with an MLU (Mean Length of Utterance) between 1.3 and 1.9 used significantly more pronouns for present referents (75%) than for absent referents (25%), indicating that their use of definite forms (pronouns in this case) correlates with the perceptual availability of the referent in the context. Similarly, Serratrice (2005) found that Italian-speaking children with an MLU of 1.5 are sensitive to disambiguation at the earliest stage—they used more overt forms like nouns and pronouns for referents with a potential competitor in the discourse or physical context, but they produced more null forms when no competitor was present.

In sum, these findings suggest that children’s situational understanding might have been eclipsed in prior experiments (Matthews et al., 2006; Campbell et al., 2000). Notably, in these experiments, young children around 2 years of age interacted with unfamiliar experimenters instead of their caregivers. Their ability for context-sensitive production may indeed have been inhibited in these experiments, given findings showing that both conversational setting and partner have an impact on how children use language (Hoff, 2010).

The claim that children are more sensitive to discourse than other properties of referring expressions becomes more questionable when we consider finer-grained properties for the referent. Studies of English-learning children’s spontaneous speech indicate that even 2-year-olds adjust their referring expressions based on different dimensions of referent accessibility (M. E. Hughes and Allen, 2013, 2015). Importantly, Hughes and Allen (2015) found that the effects of a host of referent properties on children’s referential choices are cumulative from age 2, the earliest productive stage. Their analysis of children-caregiver speech incorporated six such properties: PERCEPTUAL AVAILABILITY, ANIMACY, PRIOR MENTION, CONTEXTUAL DISAMBIGUATION, LINGUISTIC DISAMBIGUATION, and JOINT ATTENTION. Here, aside from PRIOR

MENTION and LINGUISTIC DISAMBIGUATION, all other properties are non-discourse-related. The results show that both children and their caregivers increased their use of more informative forms, i.e., lexical over null and pronominal forms, as referents became increasingly less accessible across these dimensions. This suggests that very young children track multiple properties of referents, not just discourse ones, when choosing referring expressions.

There is further reason to believe that even before age 2, learners can use pragmatic knowledge to infer the conditioning environment of different referring expressions. By 12 to 14 months, infants track the knowledge states of interlocutors and do not assume shared knowledge, demonstrating an emerging ability to differentiate between what they know and what others know (Tomasello and Haberl, 2003). These findings suggest that sensitivity to both discourse and situation begins in infancy and may go hand in hand in shaping children’s early understanding of referential forms and their conditioning environments.

Considering 1-year-olds’ pragmatic capacity, examining referring expressions in caregiver speech is essential for several reasons. First, it provides the first model for children’s referential choices. Previous studies indicate that caregiver input influences children’s output (C.-c. Huang, 2012; Yıldız and Uzundağ, 2022), even at really young ages (Guerriero et al., 2006; Mishina-Mori, 2007). Second, examining caregiver speech allows us to determine whether the conditioning environments for referring expressions remain stable across development. This is important, because variability in input could influence how easily children generalize referential forms. Some studies suggest that developmental progress may be slower if form-function mappings are less available or consistent in input (Rozendaal and Baker, 2008). While some studies indicate consistency (C.-c. Huang, 2012), variability in input remains highly possible. Indeed, caregivers may be scaffolding (Bassano et al., 2011) and adjusting

their referential uses as children develop (Rozendaal and Baker, 2008: pp.1866), given their changing assumptions about the child’s memory and referent activation status (Gundel et al., 1993). Third, caregiver speech offers us insights into how informative the input is without feedback. Some findings suggest that the role of feedback is crucial for children’s correct use of referring expressions, in view of both naturalistic (Gundel and Johnson, 2013; Orvig et al., 2010; Skarabela, 2007) and experimental (Deutsch and Pechmann, 1982; Matthews et al., 2007; Matthews et al., 2012; Theakston, 2012) observations. 1-year-olds do not yet productively use determiners, which makes parental feedback less possible, so parental input at this stage might differ from later input in its structure. Finally, caregiver speech may display properties that differ from what we expect for normal adult-to-adult speech. It is important to analyze early child-directed speech to understand what cues are available for learners to infer the meaning of definite descriptions.

Despite these insights, existing corpus studies do not extensively examine the use of definite descriptions in early caregiver input. Most studies focus on the use of lexical nouns vs. pronouns in toddlers’ input, with the only study on English determiners (Rozendaal and Baker, 2008) grouping definite and demonstrative forms together. For studies that did focus on determiners (Sneed, 2007; Fisher and Tokura, 1995), they mostly highlight prior mention as a source of anaphoric cues based on collapsed data, but it is less clear whether anaphoric cues remain consistent across different scenarios. Moreover, while most studies analyze child-directed input from age 2 onward—when a full set of referring expressions emerges—less is known about caregiver input between age 1 and age 2. Understanding input during this earlier period is crucial, as 19-month-olds already show different expectations for interlocutor interactions based on the speaker’s use of definite vs. indefinite descriptions (Choi et al., 2018).

Together, these findings suggest that children’s sensitivity to referential accessi-

bility is multifaceted, shaped by both discourse and situational knowledge. While experimental studies indicate early challenge in integrating situational cues, natural speech data suggest that children track multiple factors in producing referential forms from early on. Investigating caregiver input during the second year of life will be key to understanding how the meaning of *the*-definites is learnable, given learners' sensitivities to different sources of conditioning cues.

5.1.2 Recognizing presuppositions

Presuppositions are assumptions that a speaker takes for granted as shared knowledge between conversational partners, forming part of the background against which utterances are interpreted (Stalnaker, 1974). These assumptions are not stated explicitly but are instead signaled by *presupposition triggers* associated with specific lexical items or structures (see Levinson, 1983 for an overview). For instance, in (35), the change-of-state verb *stop* presupposes that Joe used to drive.

(35) Joe has stopped driving.

Understanding such presuppositions requires not only sensitivity to the lexical properties of individual triggers but also the ability to interpret those triggers within a broader conversational context.

Presupposition triggers are typically classified as hard vs. soft, depending on whether their presuppositions can be easily canceled or suspended (Abusch, 2002). Hard triggers, such *it*-clefts (*it was Joe who smoked*) and focus-sensitive particles (e.g., *too*), carry presuppositions that are hard to suspend. In contrast, soft triggers, like change-of-state verbs (*stop*) and factive verbs (*know*), allow for presuppositions to be canceled (Abusch, 2002; Simons, 2013) or accommodated (D. Lewis, 1979; Heim,

1982).

For language learners, particularly children, soft presupposition triggers present a notable learning challenge. This difficulty arises because the presuppositional meanings of these triggers are not always clearly signaled by surface-level discourse cues. For instance, Dudley (2017) analyzed how belief verbs *know* and *think* are used in child-directed speech. These two verbs differ in an important dimension: *know* is veridical and factive, meaning it presupposes the truth of its complement, while *think* is not. One way learners might infer the presupposition of *know* is by noticing that it tends to take declarative complements that express old or previously established information. However, Dudley’s findings show that this pattern is not consistently observable in the input. Specifically, *know* is used more frequently when its complement introduces new information (39%) than when it refers to old information (15%). As a result, learners may not receive enough clear evidence that *know* is associated with established facts. In this light, discourse cues alone may be insufficient for learners to reliably acquire the presuppositional properties of soft triggers.

Definite descriptions featuring the determiner *the* presents a complex case for learners to recognize the presuppositions they project. While *the* is traditionally considered to be a hard trigger with respect to its existential presupposition—that the referent exists (Abbott, 2006), it also behaves as a soft trigger in terms of its uniqueness presupposition (Walker, 2012). Moreover, its familiarity presupposition can be satisfied in different ways: either by referring back to entities previously mentioned (discourse-old), or by appealing to those assumed to be familiar within the situation (contextually given).

Differentiating different uses of definites is empirically necessary for a learner, as they can lead to different forms in a language. In German, for instance, the definite determiner can take different forms when following a preposition *zu* ‘to’, depending

on whether the referent has a prior discourse antecedent or is merely situationally familiar (Schwarz, 2013). Consider the two sentences below.

- (36) Hans ging zu dem Haus.
Hans went to the_{strong} house
'Hans went to the house.'

- (37) Hans ging zum Haus.
Hans went to-the_{weak} house
'Hans went to the house.'

Specifically, contraction between the preposition *zu* and the definite determiner *dem* takes place only in cases like (37), when the definite determiner expresses weak definiteness conditioned on situational familiarity. This dedicated form for encoding situational familiarity in German suggests that recognizing different uses of definites is potentially crucial for learners to acquire whether their language adopts separate forms for definiteness conditioned by discourse and situation.

This chapter investigates what cues are available in early input for toddlers to recognize the familiarity presupposition of *the*. Specifically, we ask whether easily noticeable patterns of noun repetition found in anaphoric uses of *the* are sufficient for toddlers to learn the meaning of *the*. We also examine whether toddlers can infer the presupposition of *the* through situational uses of *the*-expressions, namely for referents that are mutually familiar upon first mention.

By tracing learners' input early on and output shortly after, we aim to shed light on broader questions about how they come to extrapolate the presuppositions associated with definite descriptions in the face of unclear input. This work ultimately contributes to our understanding of how learners may develop the ability to navigate between the surface and hidden/presupposed messages of utterances—a crucial step in achieving mature communicative competence.

5.2 Study 7

This study investigates whether the surface patterns of noun repetition found in anaphoric uses are sufficient for young children to learn the meaning of *the*. The anaphoricity-driven hypothesis proposes that learners can reliably infer *the*-meaning simply through tracking distributional patterns and noticing how *the* often appears with nouns they have already heard (*a ball...the ball*). In this view, a usage preference for *the* over *a* for discourse-old referents can support learners in identifying the familiarity presupposition. According to previous research (Sneed, 2007; Fisher and Tokura, 1995), *the* is used more frequently for subsequent mentions (discourse-old referents) than *a* is, suggesting that these anaphoric cues are available in some conversational scenarios.

In contrast, the situation-driven hypothesis posits that learners must use situational understanding to infer *the*-meaning. According to this hypothesis, learners keep track of contexts/situations, and hearer-old uses of *the* can serve as situational cues for learners to identify the link between contextually given referents and *the*-expressions used for these first-mention referents, thereby inferring the familiarity presupposition of *the*-expressions. In this sense, these hearer-old uses of *the*—in which a referent is new to the discourse but presumably known to interlocutors based on shared context—could serve as an early source of evidence for the familiarity presupposition of *the*-expressions.

To evaluate these hypotheses, this study addresses two key questions, drawing on the input directed at 14-month-olds. First, to assess the anaphoricity-driven hypothesis, I examine how *the*-expressions are used depending on the discourse status of the referent—that is, whether the referent has been previously mentioned. I analyze the reliability of such anaphoric cues across different corpora. Second, to examine

Child	Age	Sex
01	1;01;14	male
02	1;01;19	female
03	1;02;01	male
04	1;01;26	female
06	1;02;11	female
07	1;02;05	male
10	1;02;04	female
13	1;02;01	male
16	1;02;03	female
19	1;02;15	male
23	1;02;09	male
25	1;02;13	male

Table 5.2: Basic information of children selected from the NewEngland corpus in Study 7

the situation-driven hypothesis, I investigate how often *the* is used for first-mention but situationally familiar (hearer-old) referents. Together, this study aims to provide insight into what types of cues are available and consistent early on in linguistic input.

5.2.1 Method

Study 7 is based on the New England corpus on CHILDES (Snow et al., 1996). It is a videotaped corpus containing longitudinal data on 52 children, with each child-parent dyad brought to the lab at three ages: at 14 months, 20 months, and between the ages of 27 and 32 months. For Study 7, we focus on corpora of 12 children (7 male and 5 female) at 14 months (Age range: 1;01;14-1;02;15). Table 5.2 summarizes information about these children in our corpus.

Basic metrics. For each corpus based on one child-parent dyad, utterances containing referring expressions are annotated in terms of the line number (in the

browsable transcripts at the TalkBank Browser), the referring expression, the referent, number (singular vs. plural), the determiner, and the sentence. The annotation of referring expressions excludes first, second, third person subjects, names, noun phrases referring to body parts, as is done in prior studies (M. E. Hughes and Allen, 2013, 2015). For determiners, we distinguish pronominal forms and non-pronominal forms. For instance, a pronominal demonstrative *that* is coded as “that-pron”, whereas a non-pronominal demonstrative *that* as in *that crayon* is coded as “that”. This will allow us to focus on non-pronominal noun phrases (e.g., *that crayon*) for our analyses, assuming that these phrases specify the property of the intended referent and can be easily understood and tracked by toddlers, following the anaphoricity-driven approach.

Discourse metrics. I use mention time as an ordinal metric to locate a particular mention within the current discourse about the same referent. To be consistent with prior research, I also label each mention alternatively as either a first mention (i.e., mention = 1) or a subsequent mention (i.e., mention > 1). For a particular referent, first mention marks its discourse property as new, whereas subsequent mention marks it as old or given.

Additionally, I use referential chain as a metric for segmenting a given corpus into separate discourse units or chains about a particular referent. Once a new referent, e.g., a football, is introduced into the discourse, the first referential chain about this referent (chain = 1, referent = FOOTBALL) begins. This chain would end if the football is not mentioned after 20 utterances since its last mention (see Guerriero et al., 2006 for the choice of 20 utterances as the discourse cutoff).² However, a new referential chain about this football (chain = 2, referent = FOOTBALL) will begin, if the football

²While some other studies (M. E. Hughes and Allen, 2013, 2015) use 5 utterances as the cut-off, it has been shown that the difference between 5 vs. 20 utterances as the cut-off is minimal (cf. Skarabela and Allen, 2002).

is mentioned a while later, being more than 20 utterances away from its last mention in the prior chain (chain = 1, referent = FOOTBALL). All referential chains about the same referent are grouped or tracked by their value under the basic metric of referent.

Usage metrics. The usage type of a referring expression is also annotated if the expression features a determiner that is either definite *the* or indefinite *a/an*.

For *the*-definites, we group the usages into four categories, based on prior research on adult corpus (Poesio and Vieira, 1997; Prince, 1992; Hawkins, 1978): anaphoric, inferable, hearer-old, and unfamiliar. The examples below illustrate these usages.

(38) Usages of *the*-definites

(a) **Anaphoric**

MOT: It's a ball.

CHI: 0.

MOT: Can you throw me *the ball*?

(NewEngland/14/03)

(b) **Inferable**

MOT: You hafta put it [=clown] back in the box.

CHI: 0.

MOT: Close *the lid*.

(NewEngland/14/07)

(c) **Hearer-old**

MOT: See *the camera*? (camera is visually present)

CHI: 0.

MOT: Taking pictures of the boy.

(NewEngland/14/13)

(d) **Unfamiliar**

MOT: Close it [=box] up.

MOT: And we'll put it [=box] over here [= on table].

CHI: 0.

MOT: Now.

CHI: 0.

MOT: Let's see what's in *the next box*.

(NewEngland/14/25)

Anaphoric uses of *the*-definites feature prior mention(s) of the referent within 20 utterances. Inferable uses of *the*-definites, also known as *bridging* (Clark, 1977) or *associative anaphoric* (Hawkins, 1978), involve reference to entities whose existence are inferable through relations such as part-whole relation (e.g., *a car...the steering wheel*) or producer relation (e.g., *a book...the author*). Hearer-old uses refer to cases where the referent is assumed to be familiar given the situation or world knowledge. As for unfamiliar uses, they refer to cases where some modifiers require the use of *the*, such as *the first box*, even upon mentioning a discourse-new referent.

For *a*-indefinites, we group the usages into four categories: hearer-new, predicative, non-specific, and generic. The examples below illustrate these usages.

(39) Usages of *a*-indefinites

(a) **Hearer-new**

MOT: Wanna read it [=book]?

MOT: Meow. *A kitten* likes to play.

(NewEngland/14/10)

(b) **Predicative**

MOT: What's that?

CHI: Kitty.

MOT: Kitty?

MOT: That's right.

MOT: That's *a kitty*.

(NewEngland/14/01)

(c) **Non-specific**

MOT: What's that [=crayon]?

CHI: 0.

MOT: Do you wanna draw on the paper?

CHI: 0.

MOT: Wanna make *a picture*?

(NewEngland/14/16)

(d) **Generic**

MOT: What's this?

CHI: 0.

MOT: Doggy.

CHI: yyy.

MOT: What does *a doggy* say?

(NewEngland/14/01)

Together, these usage metrics will allow us to examine for the first time the similarities and differences in using *the*-definites and *a*-indefinites across caregivers.

5.2.2 Results

To explore how singular definite descriptions are used in early linguistic input, we first compare their frequency to that of plural definites. In our 14-month corpora,

singular definites are far more frequent (1,574 instances) than plural definites (142 instances). This substantial difference suggests that learners are more likely to encounter singular definites early on, potentially making it easier for them to attend to and extract the uniqueness presupposition typically associated with definite determiners. While plural definites are also important for acquiring the full semantics of *the*, their relatively lower occurrence at this stage may contribute later to children’s understanding of definites. Nonetheless, a solid initial grasp of singular definites may lay the groundwork for understanding plural definites, which involve the presupposition of a uniquely identifiable plurality within the relevant domain.

We then evaluate the viability of the anaphoricity-driven hypothesis as a pathway for learners to infer the familiarity presupposition of singular *the*-expressions. We analyzed three key patterns: (1) the proportion of *the* used in first vs. subsequent mentions of referents, (2) the relative use of *the* vs. *a* for referents that are new vs. old to the discourse, and (3) the distribution of *the* and *a* across referential chains. These analyses help us assess whether early input supports the inference of familiarity based on the anaphoric use of *the*.

We first look at first mentions and subsequent mentions. The choice of first mention and subsequent mention is in line with the metrics used in prior research (Fraurud, 1990). We find that surface anaphoric cues are weakened by first-mention/situational uses of *the*, as shown in Figure 5.1. Caregivers use *the* at equal rates for first/new (51.6%) and subsequent/old (48.4%) mentions of referents, $t(22) = .76$, $p = .46$. This suggests that the discourse status of referents labeled by *the*-definites alone does not reveal caregivers’ preference for using *the* for old referents.

Next, we look at the relative proportion of *the* vs. *a* when the discourse status of the referent is new vs. old. We plot the collapsed result in Figure 5.2. Overall, we see that anaphoric cues—a preference for *the* with old over new referents—only

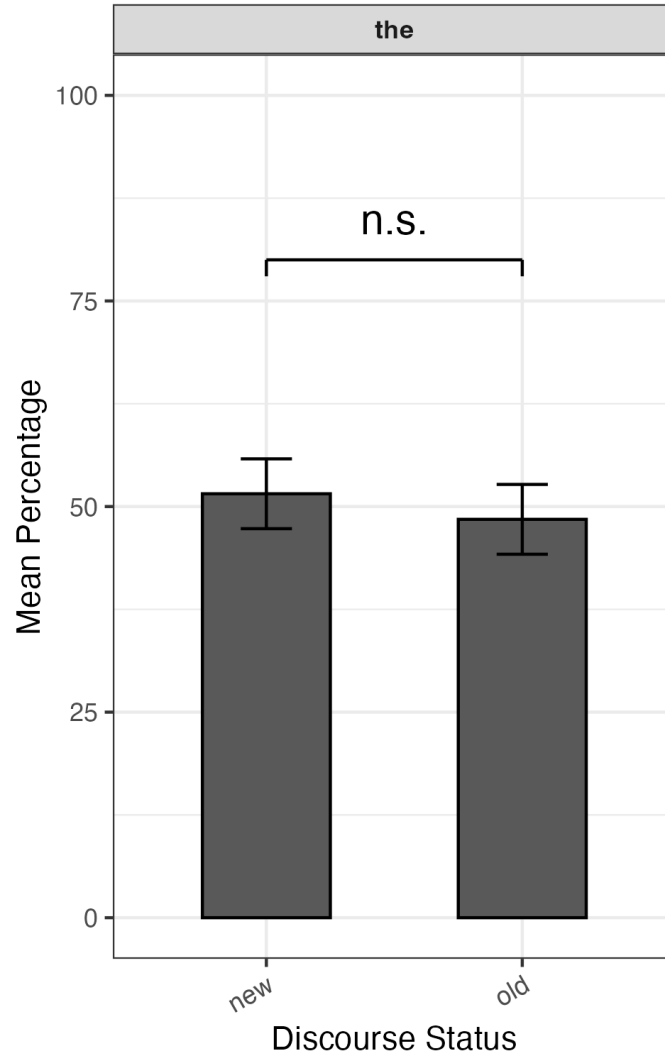


Figure 5.1: Proportion of use in first vs. subsequent mentions for ‘the’

emerge when *a* and *the* are contrasted. A paired-samples t-test revealed a significant difference between the relative percentages of *the* used for new and old referents, $t(11) = 4.6$, $p = .0008$. This indicates that learners need to consider *a*-indefinites as an alternative to identify the discourse-wise distinction.

However, this discourse contrast is not always clear-cut across scenarios. When we look into different child-parent dyads, we do observe that such anaphoric cues may not be consistently present. As shown in Figure 5.3, in 3 (A1, A7, A16) out

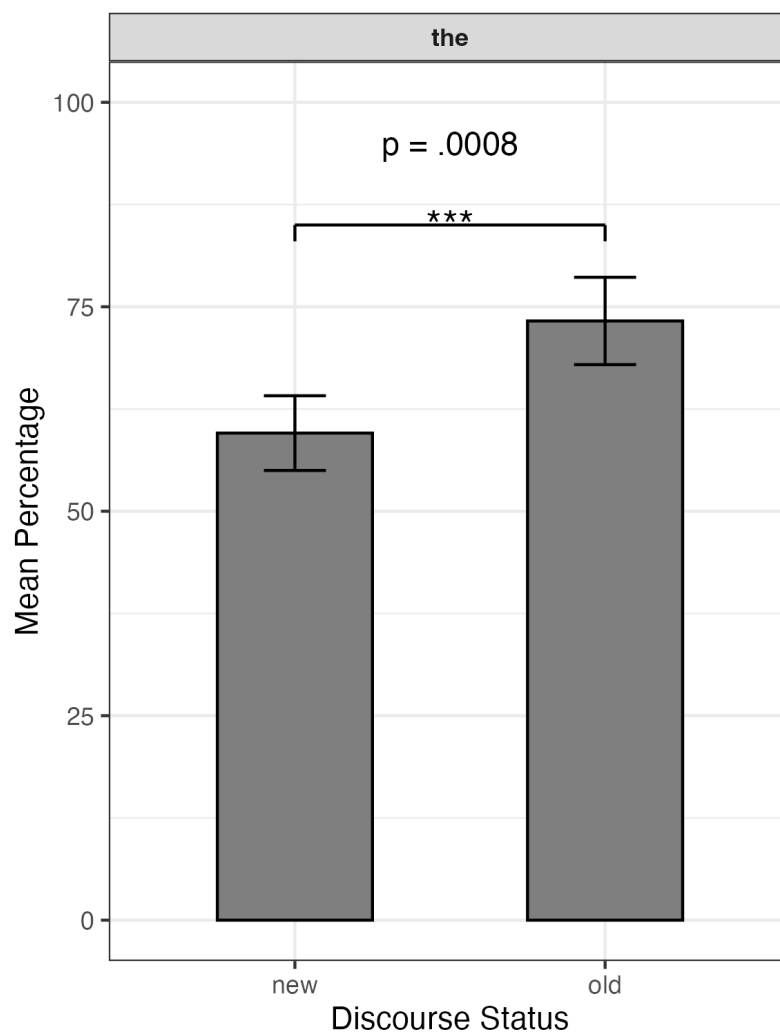


Figure 5.2: Proportion of ‘the’ vs. ‘a’ by Discourse Status of the referent

of 12 corpora, the proportion of *the* is not clearly higher when the referent is old as opposed to new.

It is possible that this degraded signal may be due to caregivers’ widespread use of *a*-indefinites in labeling (e.g., *It’s a bunny*) in subsequent mentions, as shown in Figure 5.4. This is possible, given that caregivers may repeat indefinites (e.g., *A ball! It’s a ball! Yeah, it’s a ball!*) to reinforce their children’s memory of the word label itself. Filtering out these labeling cases, however, does not sharpen the signal,

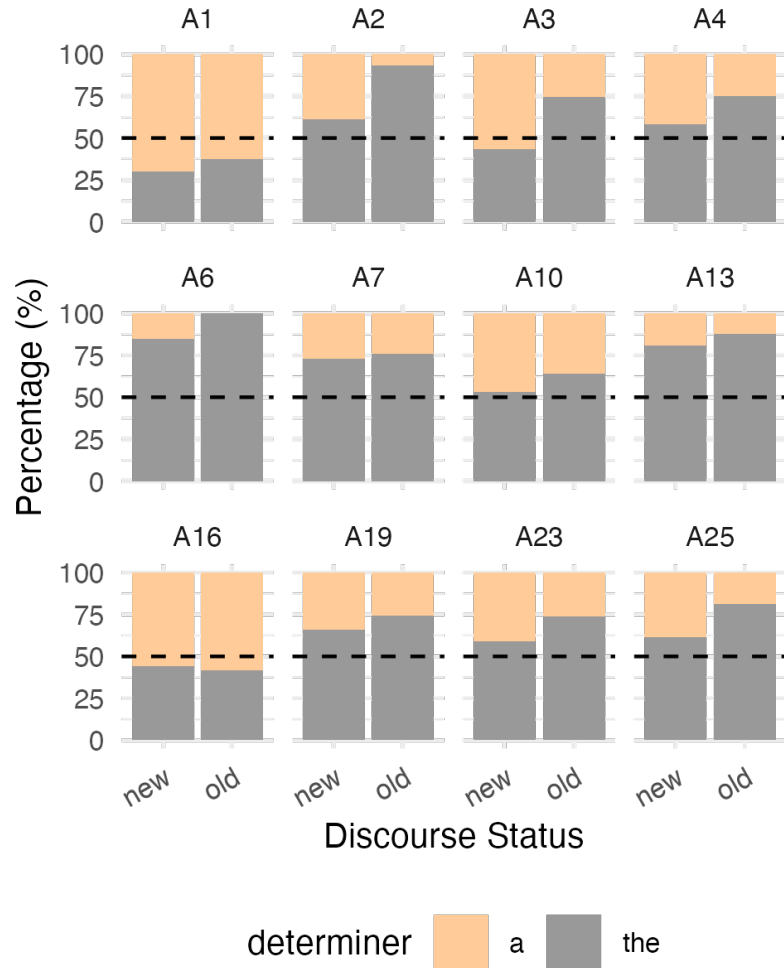


Figure 5.3: Proportion of ‘the’ vs. ‘a’ by Discourse Status of the referent across child-parent dyads

as seen in Figure 5.5. These results suggest that anaphoric cues in caregiver speech may be less clear-cut occasionally.

In addition to these collapsed cues, our results also reveal a clean pattern of incremental use of *the*-definites across referential chains. When data from all corpora are combined, we observe that the proportion of *the*-usages increases from chain1 to chain2, whereas that of *a*-usages decreases, as illustrated in Figure 5.6. This shift suggests that as discourse progresses, speakers tend to mark referents with *the*,

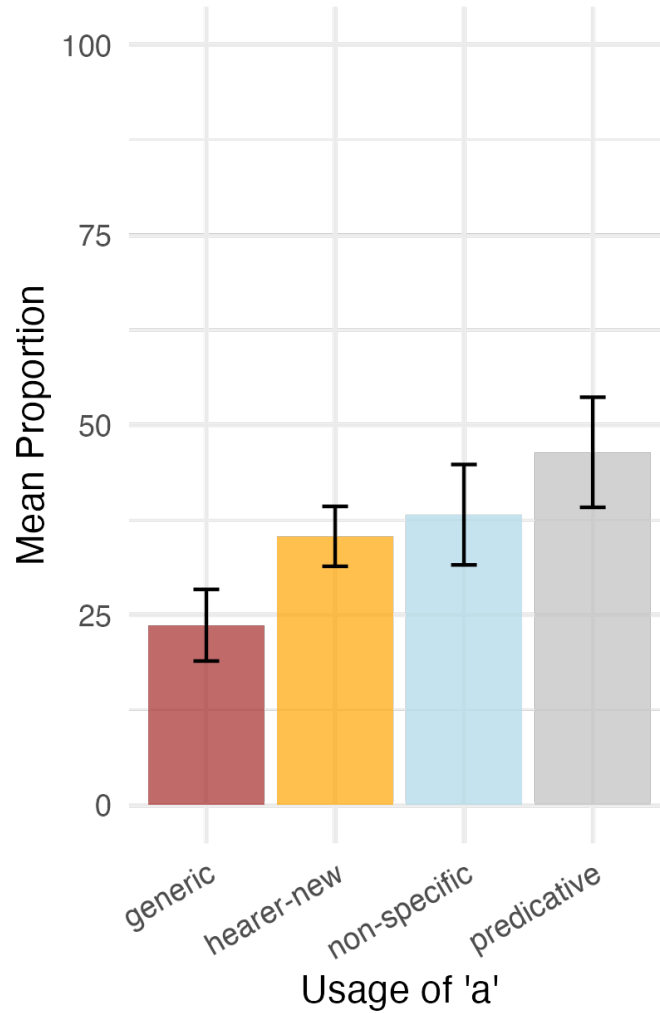


Figure 5.4: Proportion of ‘the’ vs. ‘a’ by Discourse Status of the referent across child-parent dyads

assuming shared knowledge. Crucially, this same incremental use of *the* holds across dyads, as illustrated in Figure 5.7.

However, our findings also point to another glaring problem for young learners—the overwhelmingly common use of *the*-definites for discourse-new referents. That is, in 11 (91.7%) of the 12 caregiver-child dyads, caregivers use *the*-definites as an equally or even more common alternative for *a*-indefinites when referring to entities new to the discourse, as illustrated previously in Figure 5.5. Likewise, as

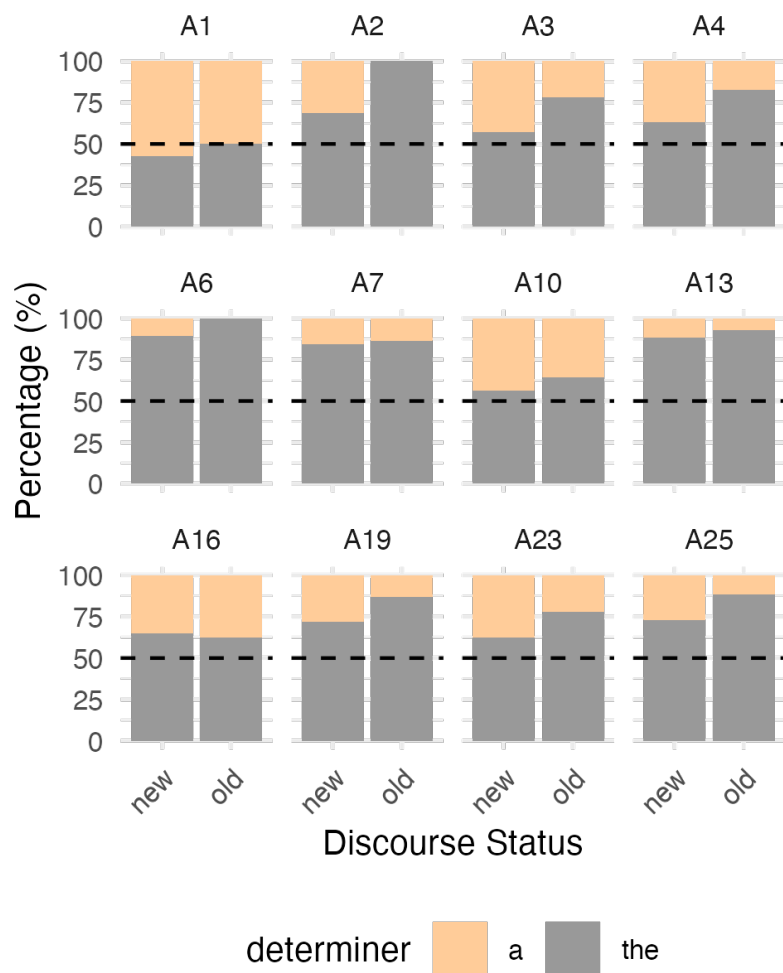


Figure 5.5: Proportion of ‘the’ vs. ‘a’ by Discourse Status of the referent (filtering predicative uses) across child-parent dyads

shown in Figure 5.1, *the*-definites are used equally frequently in first vs. subsequent mentions, a pattern that has also been observed in prior research (Sneed, 2007). These findings suggest that anaphoricity-driven learners could not simply ignore the frequent use of *the*-definites for referents new to the discourse.

We next examine the viability of inferring the familiarity presupposition of *the*-expressions through situational cues. To examine the availability of situational cues, we look at the proportion of hearer-old usages of *the*. These cases are annotated as

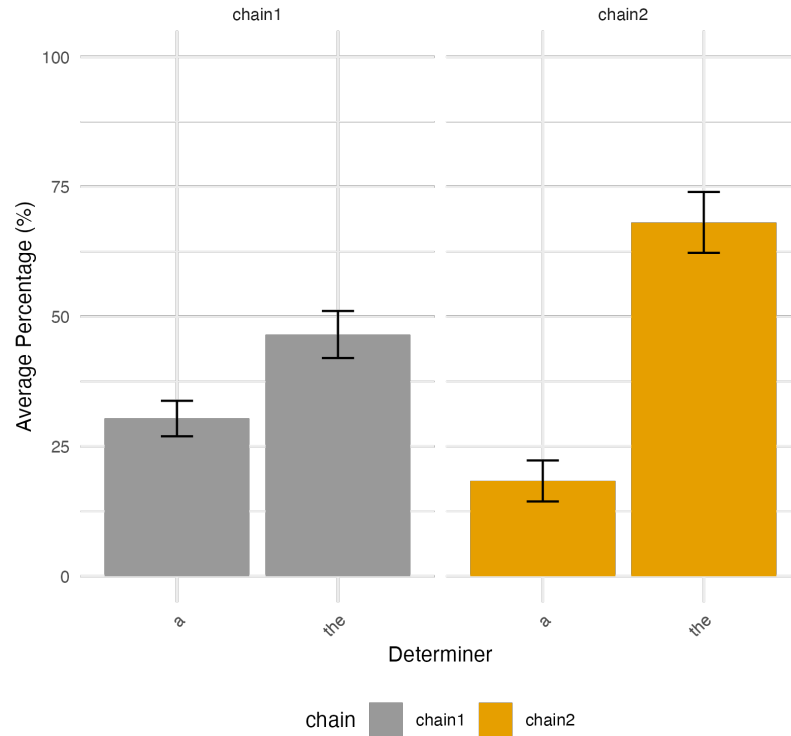


Figure 5.6: Proportion of ‘the’ vs. ‘a’ in chain1 vs. chain2

hearer-old, given that the intended referents are uniquely identifiable in the situations present in the videotaped corpora. If learners at this age are aware of the situational familiarity or identifiability of the referents, they may use these situational cues to infer the target meaning—via establishing a link between the determiner *the* and the uniquely identifiable referent it picks out given the situation.

Hearer-old uses of *the*-definites feature prominently when we examine the usage profile of *the*-definites, accounting for 32.8% of *the*-usages on average, as shown in Figure 5.8.

Hearer-old uses of *the* are also widely available in the input across individual corpora. This implies that situational cues from *the*-expressions used for situationally given referents are consistent across corpora. Notably, some caregivers (i.e., A6

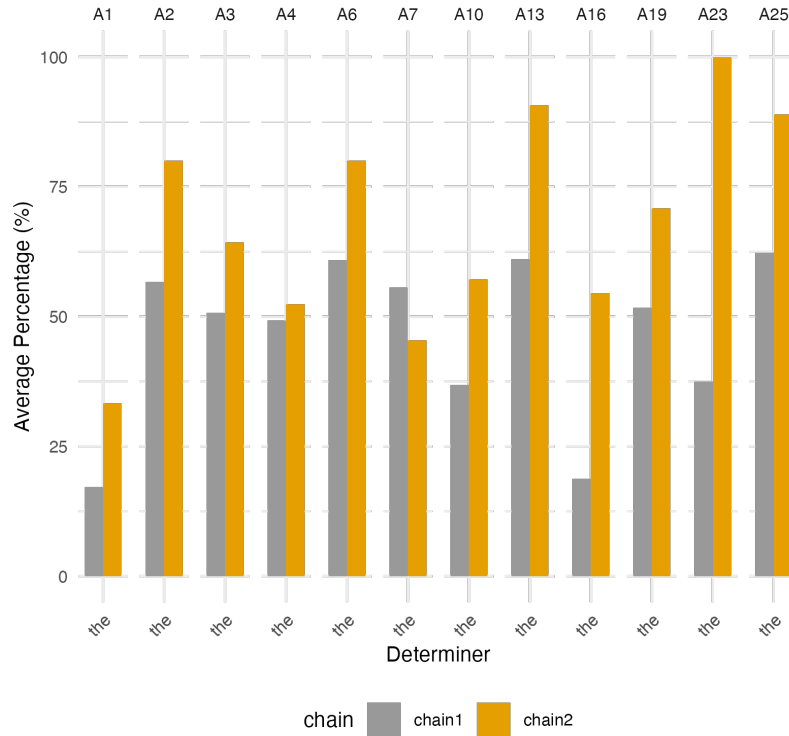


Figure 5.7: Proportion of ‘the’ in chain1 vs. chain2

and A16) even use *the* more frequently for hearer-old than anaphorically identified referents, as shown in Figure 5.9.

These findings confirm that *the*-usages for discourse-new, hearer-old referents are common in the input to 14-month-olds. In naturalistic speech, caregivers do not shy away from using *the*-definites for situationally granted but previously unmentioned referents. This means that while the familiarity of intended referents can be scaffolded through discourse, this is not always the case in early input. When faced with common hearer-old uses of *the*, learners will need to identify the intended referents given their situational knowledge.

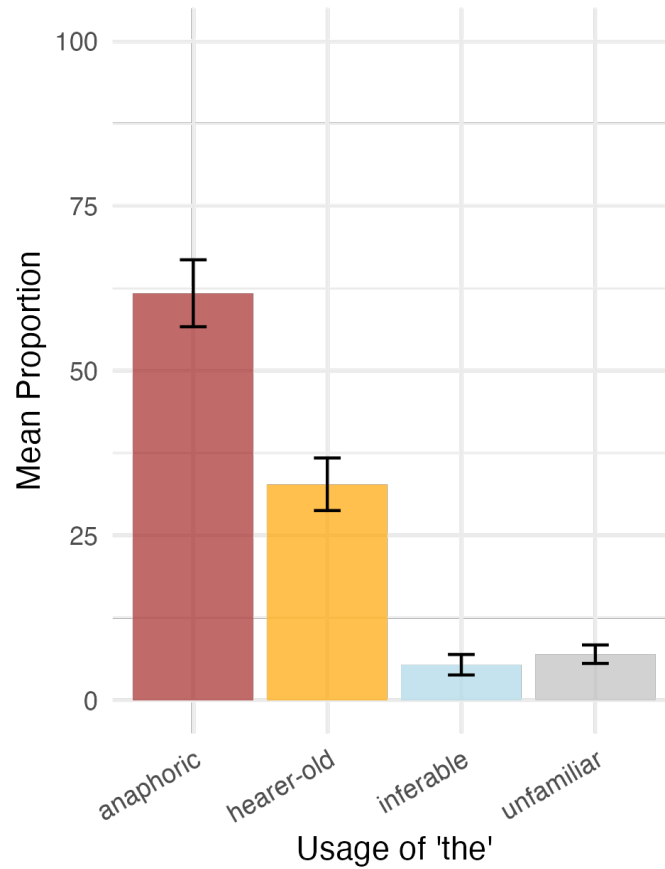


Figure 5.8: Average proportion of ‘the’ usages in chain1

5.2.3 Discussion

In Study 7, we examined caregivers’ input to 14-month-olds and assessed the availability of anaphoric and situational cues for learners’ inference of the familiarity presupposition of *the*-expressions. Anaphoric cues require learners to go beyond numerical frequency of determiners themselves and compare the relative proportion of *the* vs. *a*, and can sometimes be absent in the input. In contrast, situational cues are widely available to learners, as long as they consider shared knowledge informed by the situation in inferring the familiarity presupposition of *the*-expressions. Therefore, in terms of the availability of cues, the situation-driven inference of *the*-meaning has a

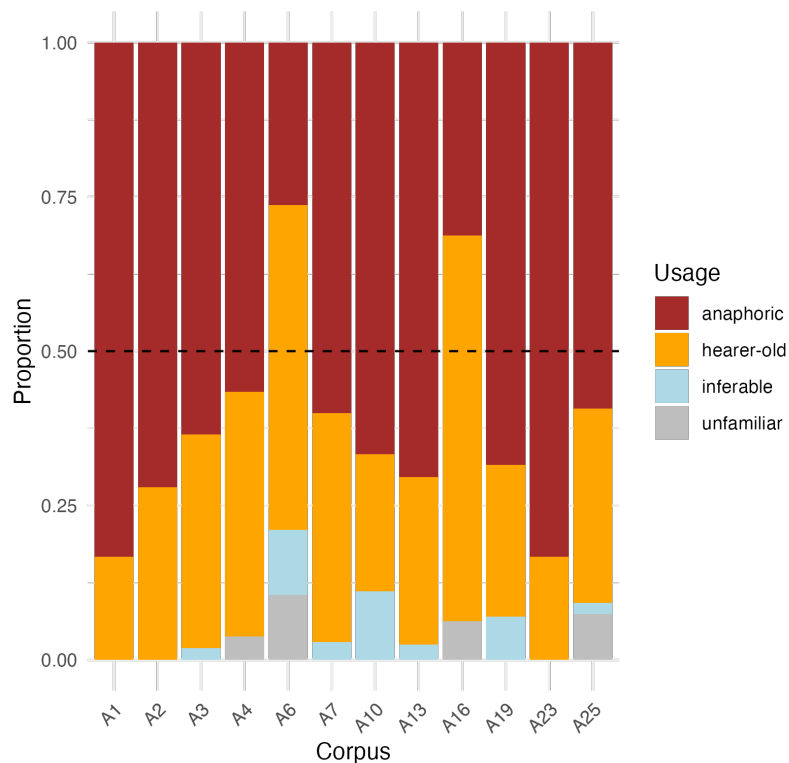


Figure 5.9: Proportion of ‘the’ usages in chain1 across dyads

slight advantage over the anaphoricity-driven approach, as situational cues are more consistently available in the input.

While anaphoricity-driven inference can in principle support learners to identify the familiarity presupposition of *the*-expressions, they can be challenging in practice, considering their usability.

On the one hand, in tracking anaphoric cues within a referential chain, learners would need sufficient attention span to relate a sequence of referring expressions to the same intended referent. This can be challenging, given that learners around one year of age tend to have a short sustained attention even with a single object in sight during free play, with a mean duration per episode of around 6.21 seconds (cf. Ruff and Lawson, 1990). This short sustained attention can take a further hit and drop

to a mean duration per episode of around 3.33 seconds given other objects in the context. In view of limitations to learners' attention span, it is less clear whether their short attention span allows for capturing relevant anaphoric cues.

On the other hand, to track anaphoric cues from a cross-chain increase in the proportion of *the*-usages, learners must actively maintain discourse referents in their mental file (cf. Recanati, 2012; Heim, 1982) over time. This is likely, as communicative acts between young learners and their caregivers may support them in maintaining and updating the representations of entities within a broader discourse file, even when these entities are displaced or lack clear spatial anchoring (Brody and Csibra, 2025; Csibra, 2010). However, there is counter evidence from comprehension studies suggesting that learners may fail to capture anaphoric uses of determiners like *the* (Modyanova and Wexler, 2007) and *that* (Ahn and Arunachalam, 2020), even up to age three. In this light, it is less clear how much these distant anaphoric cues can contribute to young learners' inference of *the*-meaning, given the lack of evidence for how learners around this age maintain and access discourse referents in their mental files.

Furthermore, for anaphoricity-driven learners, the common use of *the* for referents new to the discourse presents a puzzle. As Study 7 has shown, *the*-usages are equally common for first-mention and subsequent-mention referents, unlike pronoun *it*, which is strongly anaphoric and used predominantly in subsequent mentions. In some sense, this use of *the* for discourse-new referents can be hard to reconcile with caregivers' increased preference for using *the* as opposed to *a* for discourse-old referents.

If learners infer *the*-meaning primarily through anaphoric uses, one possibility is that these learners, puzzled by *the*-usages for first-mention referents, may mimic the distribution of *the*-expressions in their input in early production and use *the*-expressions at a similar rate for first-mention and subsequent-mention referents. This

predicts that learners will more likely misuse *the*-expressions for first-mention referents.

Another possibility is that these learners may infer a non-target meaning, positing that *the*-expressions are used for speaker-specific referents (Schaeffer and Matthewson, 2005), not necessarily ones that are mutual familiar. This is highly possible, given that it is compatible with the input that learners receive. That is, whenever *the*-expressions are used, the intended referents, by virtue of being mutually familiar, are also specific to the speaker. If learners endorse this non-target meaning, we would expect to see them consistently using *the*-expressions for referents that are specific to themselves but not known to the listener.

Nevertheless, *the*-usages for discourse-new referents, the bane to anaphoricity-driven learners, can be a blessing to situation-driven learners. If learners' inference of *the*-meaning is situation-driven, capitalizing on situational cues, these usages can be easily explained by situational familiarity. For situation-driven learners, we would expect the two non-target usages outlined above for anaphoricity-driven learners to not show in their earliest production. First, learners' use of *the*-expressions for first-mention and subsequent-mention referents should not necessarily follow the patterns of their caregivers. That is, we expect them to use *the*-expressions appropriately, either when the referent is discourse-old or hearer-old, rather than simply mimicking the proportion by their caregivers. Second, these situation-driven learners should use *a*-expressions, instead of *the*-expressions, for referents that are specific only to themselves but not known to the listener. To examine whether this is the case in early production, we turn to Study 8.

5.3 Study 8

Study 8 aims to assess two follow-up questions. First, I examine whether caregivers shift to using more discourse-new, hearer-old uses of *the*-definites. An adaptation hypothesis (Bassano et al., 2011) proposes that caregivers adapt their referential choices to match their perception of the child’s developing cognitive and pragmatic abilities. In this view, an increase in caregivers’ hearer-old uses of *the* would follow naturally if caregivers assume that their children possess more developed situational and world knowledge to understand definite reference upon first mentions.

Second, I investigate whether children’s earliest production shows two types of misuse errors that more likely follow from anaphoricity-driven inference. One is that they might be puzzled by *the*-usages for first-mention referents and simply mimic the proportion of first- and subsequent-mention referents in caregivers’ speech. As caregivers tend to use *the* for discourse-new referents above chance in input to 14-month-olds, we should expect 20-month-olds to show the same pattern. Another is that they might posit that *the*-expressions are licensed as long as the referents are known to the speaker, not necessarily the listener. Consequently, we would expect them to use *the*-expressions whenever the intended referent is known to themselves but not to the listener. However, these types of misuses should not occur in learners’ early production at 20 months, if their inference is situation-driven.

5.3.1 Methods

Study 8 is based on the same New England corpus on CHILDES (Snow et al., 1996). For Study 8, we focus on corpora of the same 12 children (7 male and 5 female) at 20 months (Age range: 1;06;26-1;08;15). Table 5.3 summarizes information about these children in our corpus.

Child	Age	Sex
01	1;06;26	male
02	1;07;23	female
03	1;08;01	male
04	1;07;16	female
06	1;08;03	female
07	1;07;27	male
10	1;07;19	female
13	1;07;15	male
16	1;07;21	female
19	1;07;19	male
23	1;08;15	male
25	1;07;14	male

Table 5.3: Basic information of children selected from the NewEngland corpus in Study 8

The same basic, discourse, and usage metrics are adopted for annotating the video corpus with the same 12 20-month-olds.

5.3.2 Results

Our findings reveal a developmental trend in how caregivers use definite descriptions: they tend to produce more hearer-old uses of *the*-expressions when speaking to 20-month-olds than when interacting with the same children at 14 months. This pattern is illustrated in Figure 5.10.

A closer examination of individual caregiver-child dyads shows that this increase in hearer-old uses of *the* is not limited a small subset of families—it holds for 9 out of the 12 dyads analyzed, as shown in the shaded area in Figure 5.11, where proportion of hearer-old uses at 20 months exceeds that at 14 months. This suggests a robust and widespread shift in caregivers’ referential behavior over time.

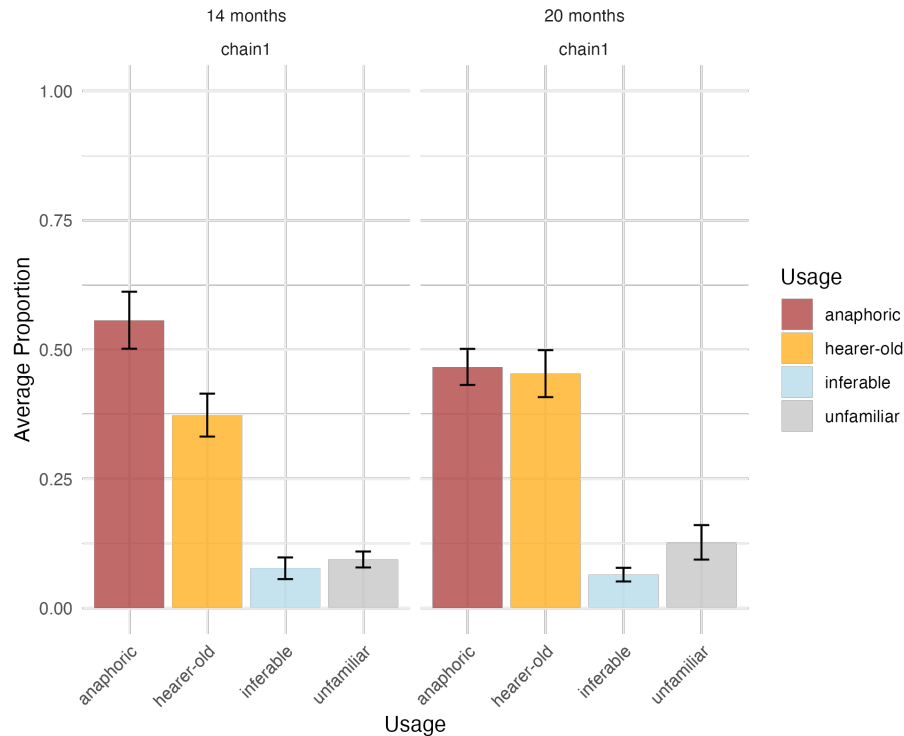


Figure 5.10: Average proportion of different ‘the’-usages at 14 months vs. 20 months

This observed increase in *the*-expressions used for referents that are already situationally familiar to the child is consistent with the adaptation hypothesis. As children grow older and become more attuned to shared knowledge and situational context, caregivers may tend to adapt their speech—using *the*-definites more often when the referent is assumed to be mutually known or contextually given. One possible interpretation is that learners’ recognition of contextually given referents may have improved substantially over the 6 months between the two lab visits (one at 14 months and another at 20 months), which might have motivated their adaptation. Nonetheless, this adaptive pattern observed across individual corpora should be treated with caution, given the size of the corpora.

Turning to learners’ output at 20 months, we first examine the determiner types that they produce in the corpora. As discussed in Study 7, *the* is used more fre-

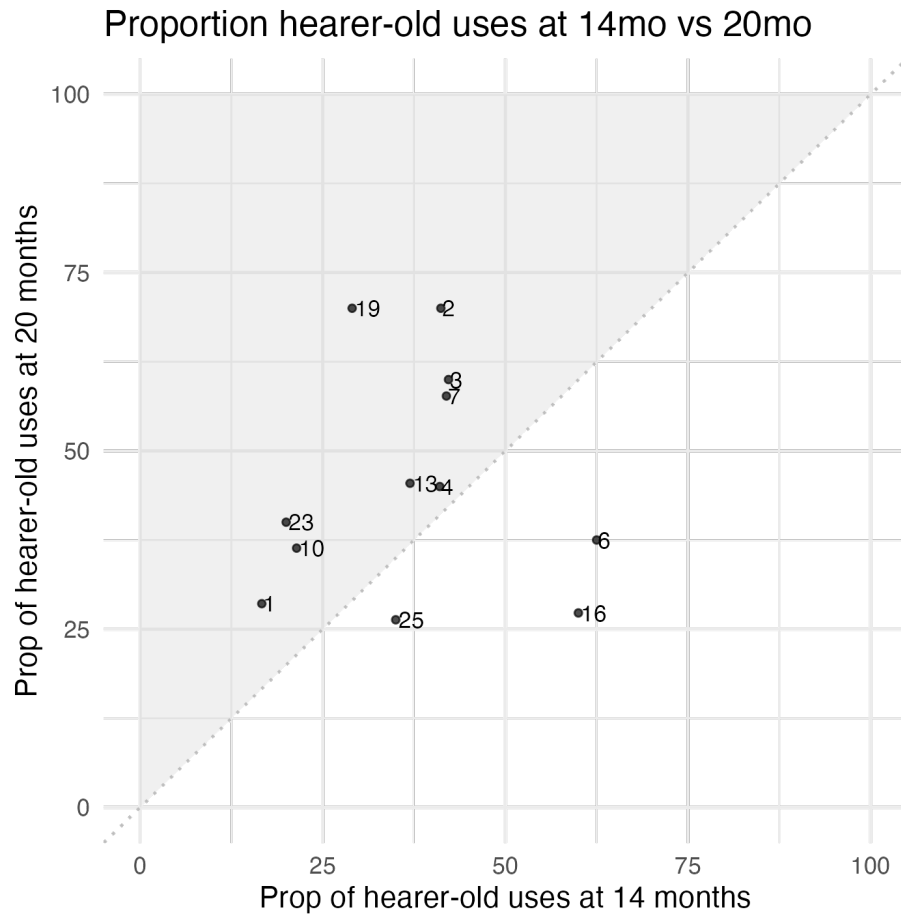


Figure 5.11: Proportion of hearer-old uses of ‘the’ at 14 months vs. 20 months

quently overall compared to *a*. If learners follow anaphoricity-driven inference and are confused by the conflicting usages of *the*-expressions for both discourse-new and discourse-old referents, they may simply mimic the same pattern by using more *the*-expressions overall in early production, even when mutual familiarity of the referent is not satisfied.

This predominant use of *the*-expressions, however, is not found in these learners’ earliest production, as illustrated in Figure 5.12. First, for toddlers who produced more than three determiners within these corpora—namely B3, B6, B7, and B10—none of them have a bias for using *the*. Second, for all uses of *the*-expressions by these

toddlers, we did not observe inappropriate uses, including those who only produced *the* but no *a* in the corpora (i.e., B13, B16, and B23). This suggests that 20-month-olds do not simply mimic their caregivers' uses of *the* proportion-wise.

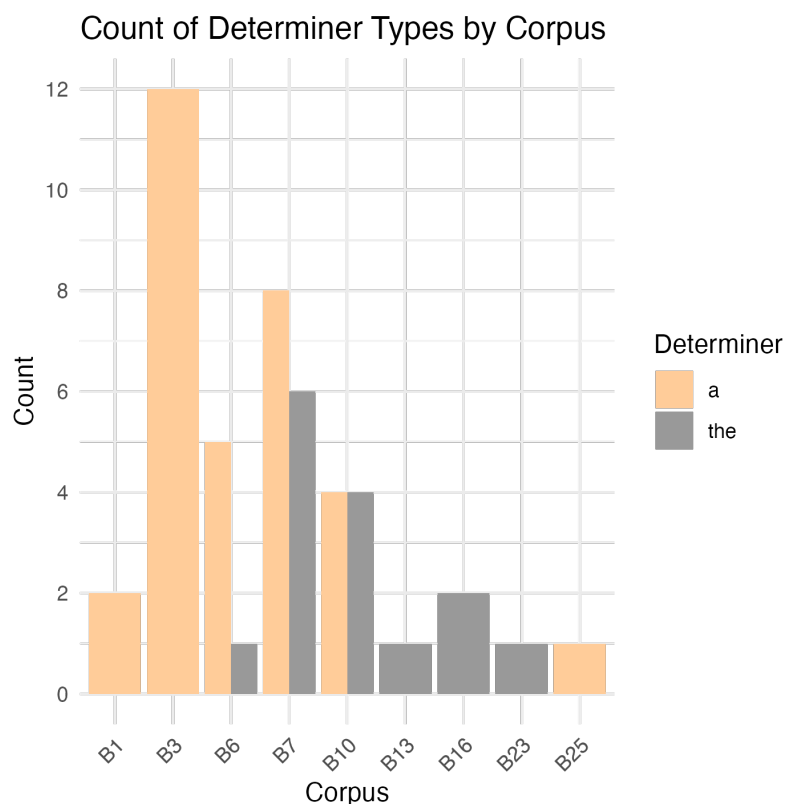


Figure 5.12: Count of determiner types (*the* vs. *a*) produced by 20-month-olds in different corpora

Our findings from the longitudinal corpus suggest that learners' determiner choices are not impacted by strong determiner bias in their caregivers, either. First, overwhelmingly more *the*-usages in learners' input does not lead to overuse of *the* in their output. Child 6, for instance, produces more *a*-indefinites (5 instances) than *the*-definites (1 instance) at 20 months. Crucially, the child produced an instance of hearer-old use of *the* with contextual support, as shown in Figure 5.13. That is, the child first pointed at a rubber duckie to draw the caregiver's attention to the in-

tended referent, thereby licensing her use of *the duck*. This appropriate use of *the* for a discourse-new referent, conversely, offers support for the situation-driven hypothesis. That is, as learners are sensitive to situational cues, they will accumulate more evidence for the target meaning of *the*-expressions when exposed to more *the*-usages.



Figure 5.13: Use of ‘the’ by child 6 at 20 months (time: 0’57’)

Second, infrequent *the*-usages for discourse-new referents in the input does not prevent a learner from acquiring the hearer-old use of *the*. This is exemplified from the following context (40) with child 16. The scene is illustrated in Figure 5.14.

- (40) Hearer-old use of *the*
MOT: You want me to turn it [= handle of jack-in-the-box]?
CHI: No. Read **the** book.
(NewEngland/20/16)



Figure 5.14: Use of ‘the’ by child 16 at 19 months (time: 2’31”)

We now examine another type of non-target usage predicted by the anaphoricity-driven hypothesis—learners might consistently use *the* for referents known only to themselves. This non-target meaning, however, is not found in learners exposed to

input with dominant *the*-usages for first-mention referents. For instance, child 7 correctly used indefinite *a* to introduce a specific referent that was discourse-new. (41) shows the discourse for the child's utterance *a box*.

- (41) MOT: Now they [= INVs] want us to play with this [= ball] for ten minutes.
MOT: But I don't know what to do with you.
MOT: You wanna play with it [= ball] some more?
CHI: No.
MOT: You sure?
CHI: **A box**.
(NewEngland/20/7)

While the mother was asking the child whether he wanted to play with the ball some more, the child switched to pointing at a specific box in a different orientation that the mother was not attending to and uttered *a box*, as shown in Figure 5.15.

Child 7's correct use of indefinite is not due to his lack of knowledge of *the*-usages. In addition to correctly using indefinite *a* for a discourse-new referent, child 7 also produced proper hearer-old (42) as well as anaphoric (43) uses of definite *the*. Figure 5.16 illustrates the scene for (42).

- (42) Hearer-old use of *the*
After the mother folded the towel they used for playing peeka-boo and put it in the box, the child bent over to reach the cover of the box.
MOT: Let's fold it [= towel] nice and neat.
CHI: Huh?
MOT: There.
CHI: Close **the box**.



Figure 5.15: Use of 'a' by child 7 at 19 months (time: 8'40")

(NewEngland/20/7)

- (43) Anaphoric use of *the*
 MOT: What's this rabbit doing?
 CHI: Ju(mp).
 MOT: Jumpin(g) rope.
 CHI: Jumpin(g) the rope.
 (NewEngland/20/7)



Figure 5.16: Use of ‘the’ by child 7 at 19 months (time: 12’33”)

We acknowledge that the anaphoric cues observed in our corpora may not fully capture the richness and variability of the input that learners receive in their everyday environments. It remains possible that outside the recorded sessions, children may be exposed to stronger anaphoric cues and can rely on them to infer the familiarity presupposition associated with *the*-expressions. However, we argue that this scenario is unlikely for two broader reasons.

First, the corpora were collected in a lab setting—an environment that was unfamiliar to both caregivers and children. This unfamiliarity creates a microcosmic

context in which caregivers are more inclined to introduce new referents and scaffold *the*-expressions through discourse (as observed in Fisher and Tokura, 1995). If anything, this lab environment likely encourages clearer discourse marking than what typically occurs at home. In contrast, everyday caregiver-child interactions take place in a highly familiar setting, where many objects and referents are shared knowledge between the child and the caregiver. In such familiar contexts, caregivers may be even more prone to use *the* for discourse-new referents, assuming shared knowledge without prior mentions. This would make it even harder—not easier—for children to rely solely on anaphoric cues to acquire the situational nuances of *the*.

Second, caregivers generally exhibit consistency in their referential style. Previous studies have demonstrated that caregiver input strongly shapes children’s output (C.-c. Huang, 2012; Yıldız and Uzundağ, 2022), even at very early stages of development (Guerriero et al., 2006; Mishina-Mori, 2007). This suggests that children accumulate experience with a relatively stable input style, such that they themselves would take on these referential habits over time. If a caregiver frequently uses *the* for discourse-new referents, it is unlikely that their style would shift dramatically in more naturalistic home settings, where more contextually given referents encourage hearer-old uses of *the*. Rather, the same referential tendencies are likely to persist, further limiting the informativeness of anaphoric cues.

In sum, the findings from Study 8 point to two key conclusions. First, on the input side, caregivers show an increase in the use of hearer-old *the*-expressions when interacting with 20-month-olds compared to 14-month-olds. Second, on the learner side, despite the variability in input profiles of different *the*-usages across caregivers, children at around 20 months show successful use of *the* in both hearer-old and anaphoric contexts. They do not mimic the distributional patterns in their caregivers’ input or arrive at a non-target meaning of *the*. This suggests that children’s developing

understanding of *the* cannot be fully explained by reliance on surface-level anaphoric cues. Instead, it likely reflects deeper generalizations drawn from both anaphoric and situational cues across multiple contexts.

5.4 General discussion

Across two studies using videotaped longitudinal corpora, we examined how early caregiver input may support young learners in acquiring the familiarity presupposition associated with *the*-expressions. Our findings show that, by 14 months, learners are exposed to a mix of both anaphoric and situational cues for *the*-usage. Crucially, even when input quality varies across scenarios—especially in terms of the availability of anaphoric cues—learners still go on to produce appropriate situationally conditioned uses of *the* at 20 months. This provides evidence in favor of the situation-driven inference of *the*-meaning.

One central insight from our analyses is that the anaphoric cues in child-directed speech for inferring the familiarity presupposition of *the*-expressions are available but not consistently so. While caregivers tend to use more *the*-expressions relative to *a*-expressions for discourse-old referents, in many other cases, they use *the* also for discourse-new referents that are situationally given—objects present in the immediate physical environment or highly salient in the conversational context. If learners were to rely solely on anaphoric cues to infer the conditions for using *the*, they might find *the*-usages for discourse-new referents confusing. As a result, they may mimic these uses by caregivers proportion-wise in early production, without satisfying their requirement for situational familiarity, or they may mistakenly conclude that *the* can be used with any specific referent the speaker has in mind, regardless of whether it is known to the listener, as this non-target meaning is perfectly compatible with their

input.

While local anaphoric cues are variable across scenarios, non-local anaphoric cues of *the* are widely available across dyads. That is, *the*-expressions are used more for mentioning established referents a while later. However, to exploit these distantly anaphoric cues, learners must maintain previously established discourse referents in their mental file to recognize their familiarity upon future mentions. This concern is well-founded, given evidence from comprehension studies suggesting that learners may fail to capture anaphoric uses of determiners (Modyanova and Wexler, 2007), especially when these determiners are not strongly anaphoric in use compared to pronouns (Ahn and Arunachalam, 2020). While successful retrieval of distantly mentioned referents cannot always be guaranteed, this may be possible as learners develop more sustained attention with time to capture extra signals in longer stretches of the discourse. We need more evidence for learners’ capacity for retrieving discourse referents before 18 months in future research.

Despite potential challenges posed by the input, our longitudinal results suggest that learners are not led astray. While previous studies indicate that caregiver input influences children’s output (C.-c. Huang, 2012; Yıldız and Uzundağ, 2022), even at really young ages (Guerriero et al., 2006; Mishina-Mori, 2007), this influence does not seem to carry over to learners’ usages of *the*. In fact, regardless of the variable quality of anaphoric cues in their input, learners around 20 months still produced *the*-expressions appropriately—using them in both situationally given and anaphoric contexts. This outcome points to the crucial role of pragmatic competence in development: to properly solve the puzzle with *the*-usages for first-mention referents, young learners can leverage situational relevance and shared knowledge to infer the link between these usages and the unique identifiability of the intended referents, given the referential contexts. While surface anaphoric patterns of noun repetition

may be mixed and lead them astray to consider a non-target meaning—*the* is used for speaker-known referents—we do not observe this in the early output of these learners.

Furthermore, learners’ burgeoning context-dependent flexibility in determiner choice shown in Study 8 is in line with experimental findings on their comprehension. It has been shown that by around 21 months, children can use situational cues to resolve ambiguous references associated with *the*-expressions. For example, in Grosse et al. (2010), toddlers understand that when a speaker, hands free, utters a superficially ambiguous request like *give me the battery* given two batteries in the situation, the speaker is asking for the battery that is far away and beyond the speaker’s reach, as opposed to one that is close to the speaker. This demonstrates that learners around this age could integrate situational context and speaker perspective into their understanding of definite descriptions. Our findings add to this picture by showing that similar context-sensitive reasoning is already emerging in learners’ spontaneous production by 20 months.

More broadly, our findings bear on the understanding of young learners’ capacity for pragmatic reasoning. Young learners’ failure in deriving scalar implicature is a typical case where researchers attribute their performance to immature pragmatic reasoning skills, in domains like disjunction (Chierchia et al., 2004), aspectuals (Papafragou, 2006), modals (Noveck, 2001), and quantifiers (Pouscoulous et al., 2007). In the domain of reference, children have also been reported to show difficulty with deriving scalar implicatures with indefinite descriptions (van Hout et al., 2010). These cases, where learners fail, represent a general process where pragmatic reasoning is involved in judging the pragmatic appropriateness of scalar statements (Papafragou, 2006). However, for the situation-driven inference about the presupposition of *the*-expressions discussed in this chapter, what is involved is a different process, where learners are presumably practicing spontaneous pragmatic reasoning. In this light,

our findings suggest that while other applications of pragmatic reasoning might still be developing in early development, spontaneously computing pragmatic information underlying the linguistic signals is what learners are capable of, even for early toddlers.

To conclude, these results highlight the power of situational reasoning in early referential development. Even in face of surface anaphoric cues that can be degraded in quality, learners appear to make principled generalizations about the use of *the*, guided not just by anaphoric patterns but by their broader understanding of situational context. Attending to both anaphoric and situational dimensions of input allows learners to avoid overgeneralization and build a nuanced understanding of definiteness from an age younger than previously assumed.

Chapter 6: Summary and outlook

This dissertation explored how young learners approach linguistic structures and meanings from limited input, focusing on two domains: questions and reference. Across both, the findings reveal that children exhibit striking pragmatic sophistication early in development.

In question comprehension, I showed that despite incomplete structural knowledge, learners by 15 months can interpret instrumental *with*-questions (*What did she break the stick with?*)—a structure sparse in infant-directed speech—demonstrating their ability to generalize beyond the input. In understanding reference, our results indicate that children produce adult-like *the*-expressions from the earliest stages, suggesting an early grasp of definiteness conditioned by discourse and situation.

Together, these findings suggest that young learners are well-prepared to acquire complex syntactic and semantic aspects of language. Whether interpreting questions involving non-local syntactic dependencies, or producing referring expressions involving adaptation to situation, they bring powerful inferential abilities that extend well beyond what they hear.

The following sections of this chapter offer a synthesis of the key findings presented in each chapter, drawing connections between them to highlight a cohesive developmental picture. I then turn to a discussion of the implications and broader implications for our understanding of early language acquisition, and consider how these findings inform both theoretical perspectives and methodological approaches to

further investigating the subtle traces of linguistic knowledge in the minds of young learners.

6.1 Summary

Chapter 2 investigated how infants at 15 months comprehend *wh*-questions, using instrument questions that involve a stranded preposition *with*, such as *What did she break the stick with?* The transitivity-driven hypothesis predicts learners to succeed, if they essentially track the selectional properties of transitive categories, including preposition *with*, in responding to argument questions. The findings lend support to this hypothesis—infants were able to successfully respond to prepositional object questions. Importantly, their success is not contingent on usage factors like the frequency of the question type or verb-specific expectations. This suggests that young learners are pragmatically sophisticated in their response to different argument questions and may bootstrap complex, non-local dependencies through simpler, local dependencies.

Chapter 3 turned to the production of referring expression by preschoolers, specifically examining their use of *the*-expressions in naturalistic contexts. Prior studies using elicited production tasks have claimed that children between age 3 to 5 overuse *the*, based on a deficient-knowledge hypothesis—the idea that children lack adult-like cognitive or linguistic competence. In contrast, the task-performance hypothesis suggests that these patterns may reflect performance issues specific to experimental tasks. Across three analyses of natural mother-child interactions, I found that preschoolers' use of definite and indefinite determiners was closely aligned with their mothers'. Their rate of miscommunication was minimal, and naïve adults were equally accurate in guessing children's and their mothers' referential choices. These findings strongly

support the task-performance hypothesis and reveal that, in natural conversations, children already demonstrate adult-like usage.

Chapter 4 further tested this idea with task performance, using an improved elicited production task that maintained the conversational nature of natural interactions while allowing for controlled variation in referential contexts. The results showed that even 3- to 4-year-olds adapt their use of referring expressions appropriately, using *the* only when the referent was uniquely identifiable within the referential context. These findings provide direct evidence that young children possess the grammatical and pragmatic competence needed for appropriate use of definite determiner *the*, and that previously observed errors in the literature likely stem from task demands rather than systematic deficits.

Chapter 5 examined cues available in the input to learners around 14 months that signal the meaning of *the*-expressions and analyzed their learning outcome at 20 months to consider what inferential processes they need to engage in to arrive at the target meaning. Anaphoric cues (*a ball...the ball*) come from the increased use of *the* over *a* for old as opposed to new referents. While available, these signals are found to be sometimes degraded and can lead learners astray, given the equally frequent *the*-usages for old and new referents. However, situational cues—from *the*-usages for contextually given referents—are consistently available across scenarios. These signals can elucidate common uses of *the* for discourse-new referents. Importantly, longitudinal analyses revealed that despite the variable quality of anaphoric cues in input to 14-month-olds, learners are not misled in their use of *the* at 20 months—they show appropriate uses of *the* in both anaphoric and situational contexts. These findings suggest that learners capitalize on situational cues rather than solely surface anaphoric ones. This provides further evidence for learners' capacity for inferring presuppositions underlying linguistic signals through situational reasoning.

6.2 Implications

This dissertation provides a detailed picture of how young learners make inferences about linguistic structures and meanings, from resolving stranded *with*-questions to producing contextually conditioned *the*-expressions. In this section, I discuss the implications from the two case studies in terms of how learners navigate syntactic dependencies and referring expressions.

6.2.1 Navigating syntactic dependencies

As shown in Study 1 and 2, 15-month-olds succeed in comprehending stranded *with*-questions like *What did she break the stick with?*, despite the sparsity of such structures in input. According to prior research, this successful comprehension of questions involving movement dependencies (Gagliardi et al., 2016; Perkins and Lidz, 2020) happens before their knowledge for movement dependencies emerges around 18 months (Perkins and Lidz, 2021).

One implication is that probing learners' comprehension of linguistic structures can inform our understanding of the growth of linguistic knowledge, at a point in development where production is still not possible. These preposition-stranded or P-stranded questions have been shown to emerge in learners' spontaneous production generally after two and a half years of age (Sugisaki and Snyder, 2002, 2006).

While learners' success with stranded *with*-questions at 15 months reported in this work does not necessarily imply their knowledge with P-stranded questions, at the very least, this points to the possibility that learning about P-stranding may happen early on in development. Crucially, this early learning may happen for good reasons. On the one hand, it helps learners constrain the space of hypotheses about grammatical movement dependencies in their language. The possibility of stranding prepo-

sitions is in fact rarely attested cross-linguistically. According to previous research (Merchant, 1999), this happens in only a small set of languages, mostly inside the Germanic family. In all other languages, the only way of displacing a *wh*-phrase after a preposition is pied-piping (Ross, 1967), namely to have the preposition displaced together with the *wh*-phrase itself. On the other hand, once learners have registered the possibility of P-stranding in their language, they will get the application of this stranding strategy in other related structures for free. Specifically, they will learn that the same *wh*-movement dependencies would be licensed in other environments, such as topicalization, relativization, and comparatives.

Moreover, our work has important implications for how learners recover structures in early sentence processing. In this work, we focus on whether the location of the locally missing content matters for how young learners recover structures. The answer is no, based on our results. The missing noun phrase in our study happens after a transitive preposition, a scenario that is much less frequent in the input compared to one after a transitive verb. Despite this difference in the location of the missing content, learners manage to interpret *with*-stranded questions. This invites future studies to look into another dimension of the missing content—the syntactic category of the content to be recovered. While our study provides supporting evidence for young learners’ general ability to recover DPs, regardless of their locations (44), it remains an interesting puzzle whether this ability to recover structures extends to a different category, namely a VP, as shown in the case of VP ellipsis in (45).

(44) What did she break the stick with [_{DP} ~~what~~]?

(45) I watched the movie this morning, and Joe did [_{VP} ~~watch the movie~~] yesterday.

This will help us answer an equally important question as to what conceptual knowl-

edge can be mobilized to support the recovery of structures. Our findings, together with those from prior research (Gagliardi et al., 2016; Perkins and Lidz, 2020), suggest that young learners can exploit their knowledge about event participants to respond to missing DPs. However, it is less clear whether they can use their knowledge about events themselves to respond to missing VPs. If the mechanism for recovering syntactic structures in early development is one that is truly structure-neutral or category-neutral, the prediction is that learners around the same age should also show correct response to (45) with a locally missing VP. However, it is also possible that learners at this age may fail with (45). One reason could be due to the different natures of the type of dependency involved in (44) and in (45)—the former is a case of syntactic dependency, whereas the latter is one of anaphoric dependency. Specifically, the former syntactic dependency is one related to displacement, meaning that the displaced phrase must surface elsewhere inside the same clause. In contrast, the latter anaphoric dependency is one that is not bound to the same clause, as shown in (46).

(46) Matthew: I watched the movie this morning.

Will: Joe did [_{VP} ~~watch the movie~~] yesterday.

It is possible that recovering structures involved in anaphoric dependencies may follow a different mechanism.

Furthermore, our findings relate to broader discussions about early learning. While early learning may be input-driven, starting from the more frequent patterns (e.g., in segmenting words with frequent determiners, Shi et al., 2006), it can also be expectation-driven, as learners learn from the more unexpected patterns (e.g., in identifying word boundaries with low transitional probability, Saffran et al., 1996,

Teinonen et al., 2009). In the domain of learning syntactic dependencies, what drives learning may differ, depending on the specific type of syntactic dependencies. While morpho-syntactic dependencies have been shown to be input-driven (Santelmann and Jusczyk, 1998; Omaki et al., 2015), movement dependencies involving locally missing structures seem to be driven primarily by local structural expectations. As our work has shown, this expectation is derived from learners’ observation of transitive categories like transitive verbs and prepositions, such that when the expected transitivity is violated locally with missing or displaced syntactic structures, learners will respond to the unexpectedly missing element. Crucially, through testing learners with stranded *with*-questions, we have demonstrated that learners’ resolution of the missing argument is not input-driven, given the underrepresented prepositional object gaps in their early input. Moreover, through using verbs that do not co-occur with instrumental *with*-phrases in declarative clauses, we also illustrate that learners’ structural expectations are not subdued to potential lexical expectations associated with verbs in particular.

In terms of how learners may acquire non-local dependencies, our findings also point to a new possibility—they may bootstrap the dependencies from the much smaller set of transitive prepositions. This differs importantly from a previous hypothesis in Perkins (2019, 2022), which suggests that learners may bootstrap non-local dependencies from the set of transitive verbs they have learned from the input. Transitive verbs can vary drastically in their distributional properties in declarative clauses, depending on their semantic properties (Levin, 1993; Levin and Hovav, 2005). In contrast, prepositions are highly consistent distribution-wise in selecting for a local complement. In this light, stranded prepositions, though less frequent compared to stranded verbs, send out a stronger signal for a locally missing argument. To understand that English allows stranded prepositions, learners can capitalize on

prepositions that tend to be stranded more frequently among others. For instance, we found that the preposition *about* is frequently stranded, for about 27.8% of the time, as in *What are you fussing about?* Cases like this may allow learners to identify the syntactic dependency between the displaced *wh*-phrase and the stranded preposition within an even narrower time window in development.

6.2.2 Navigating referring expressions

From Study 3 to Study 6, I have shown that young learners use *the*-expressions in an adult-like manner, both in natural production and in elicited production. With Study 7 and Study 8, I further demonstrated that despite the variability in how *the*-expressions are conditioned in caregivers' input to 14-month-olds, learners show context-sensitive uses of *the*-expressions around 20 months. This points to two key implications.

First, our findings allow us to better understand how learners can uncover presuppositions underlying linguistic signals from an early age. While certain discourse patterns may provide learners with an idealized environment for proposing preliminary hypotheses about the hidden rule, they may also encounter other scenarios where the discourse pattern may be unclear or conflicting with their first-pass hypotheses. That is, relying solely on discourse patterns may lead them to first arrive at non-target meanings in an unbounded space of hypothetical meanings. Contrarily, attending to situational cues from different aspects of shared knowledge can help them land safely on the target meaning.

The acquisition of the familiarity presupposition of *the*-expressions illustrates the benefits of the situation-driven path of inference. A non-target meaning of *the* in picking out speaker-known referents, as opposed to mutually known referents, is

in principle compatible with learners' input, but this could lead learners to misuse *the*-expressions in cases where mutual familiarity of the intended referent is not established—when the intended referent is known only to the speaker but not the listener. Nevertheless, the situations of utterance can support learners to identify the link between *the*-expressions used and the referents they pick out from their shared contextual knowledge. In other words, situational dynamics are a rich source for inferring target meanings—they help learners to circumvent non-target possible meanings in the enormous space of hypotheses.

Second, our findings point to the complex interaction between linguistic performance and non-linguistic demands, especially in the domain of pragmatics. For instance, when we probe the uniqueness presupposition of *the*-expressions, it is non-trivial to consider how the relevant linguistic competence interfaces with multiple fronts in cognition. How experimental methods work is by itself a science that cannot be ignored.

For instance, one common method for eliciting referring expressions from children is story-telling. In story-telling with pictures, how the pictures are presented visually in sequence (Cohn, 2020), how colors are used to differentiate characters (whether instances in two pictures are the same or co-referential; Emslie and Stevenson, 1981), and how children encode these information in language—their narrative skills (Gagarina and Bohnacker, 2022)—are all components closely tied to each other and can be consequential for our interpretation of the findings—whether they reveal young learners' deficiency in knowledge or simply performance. In this light, story-telling or story completion without pictures can be another can of worms. Whether children can update information through purely verbal narration depends heavily on their maintenance of information in memory (Maratsos, 1976) as well as their engagement in the task.

Another crucial consideration is the clarity of the referential domain. As we are testing how learners evaluate the uniqueness of the intended referent within a target domain, it is fundamental for us to create a domain that is easily trackable for young children (Trueswell et al., 2011). Once the domain is trackable, the number of objects within that domain needs to be visually transparent and easily encodable, in view of children’s number development (Starkey and Cooper Jr, 1980; Xu et al., 2005; Dehaene, 1992). In this regard, our Study 6 consistently presents preschoolers with two objects in the referential domain, a minimal number of objects needed for manipulating the uniqueness status of the target object/referent in a slightly larger domain. This differs crucially from the number of objects used in previous research (Maratsos, 1976; Schafer and de Villiers, 2000; van Hout et al., 2010). The ease of evaluating referent uniqueness may be key to why we found very few cases of misuse of *the*-expressions, unlike the high error rates observed in previous studies. Our results also invite future research to further manipulate objects in the referential domain to better understand children’s referential choice with minimally different referential contexts. For instance, in Study 6, we presented two identical boxes to elicit indefinite descriptions for one of the boxes. It would be interesting to examine if presenting one bag and two boxes would boost children’s use of definite descriptions for one of the boxes—a pattern that has been previously observed (van Hout et al., 2010).

Last but not least, audience design can be equally important for our understanding of how learners navigate referring expressions. In our task for Study 6, we created a context where the child participant was asked to help Kermit the Frog understand what he himself could not see behind the walls. Moreover, whenever Kermit asked children what they saw, he would explicitly say *I really want to know what’s behind this wall, but I can’t see it from here*. This instruction may highlight the need for the

speakers, namely children, to consider the perspective of their listener or audience.¹ That is, the inclusion of these perspective-related reminders may also support children in adapting their referring expressions for the sake of their audience.

6.3 Broader implications

This work offers three broader insights into early language acquisition. First, it reveals how learners combine both linguistic and conceptual knowledge to comprehend and produce sentences before arriving at adult-like knowledge of structures and meanings. Second, it highlights the linguistic input available to young learners and how the inferential systems, recruiting cues from local structures and contextual knowledge, enable language acquisition in the face of sparse and unclear input. Third, it contributes to experimental methodology by identifying effective ways to probe children’s underlying linguistic competence through performance.

6.3.1 Early integration: Linking language to concept

A central finding of this dissertation is that once learners grasp key concepts, they can immediately recognize the link between these concepts and the language—and use that knowledge to process and produce speech with striking efficiency. This rapid alignment between conceptual understanding and linguistic form reflects a powerful learning strategy: learners do not simply accumulate structures in isolation but actively integrate information across linguistic and conceptual domains. Here, I explain how this early integration of linguistic and conceptual knowledge—seamlessly linking meaning and form—is illustrated in both case studies presented in the dissertation.

¹In a pilot study for Study 6, the experimenter did not turn sideways when a child participant was interacting with Kermit. The result was that the child would sometimes look at the experimenter when answering Kermit’s question, and the child tended to use inappropriate *the*-expressions with referents that were not unique in the referential context.

In question comprehension, learners as young as 15 months of age can respond correctly to argument gaps that follow an instrument-introducing preposition *with*. Their success can be better contextualized in light of the knowledge they already possess at this age—their conceptual understanding of events and their sensitivity to local structural cues in language. By 11 months, they already represent instruments as integral event participants (He, 2015), and just a few months later, at around 16 months, they know that the complement of *with* bears an instrument role (Lidz et al., 2017). As soon as learners show sensitivity to the instrument-introducing preposition *with*, they can actively use their event knowledge to seek for unmentioned instruments in stranded *with*-questions. This suggests that once the conceptual groundwork is in place, learners can latch onto partial linguistic cues for missing arguments and find a way to satisfy them with their conceptual correspondences.

In referential production, learners' ability to incorporate listener perspectives into their own use of determiners emerges shortly after the presence of their embryonic perspective-taking skills. Prior studies show that learners consider others' experience in identifying their intended referents between 12 to 18 months (Tomasello and Haberl, 2003). This dissertation shows that by 20 months, learners can correctly produce *the*-expressions conditioned by discourse, i.e., with mentioned referents, as well as ones conditioned by context/situation, i.e., with mutually known but not explicitly introduced referents. Similarly, by 3 years of age, right when learners exhibit emerging sensitivity to act on others' perspectives in social interactions (Brezack et al., 2021), they show more sophisticated proficiency in adapting their referring expressions for listeners in elicited production, depending on whether the intended referent is uniquely identifiable in the referential context.

In sum, across both question comprehension and referential production, young learners display a striking ability to link abstract conceptual understanding—thematic

roles and shared knowledge—to surface linguistic forms. These findings suggest that once learners’ core conceptual and social foundations are in place, language acquisition can proceed efficiently, as they begin to map familiar meanings onto linguistic structures.

This pattern of learning is not isolated to the domains examined in this dissertation. It resonates with findings from other domains of language development, such as verb learning. In these cases, learners can use the distributional patterns of verb within sentences to predict their likely meanings (Gleitman, 1990; Fisher et al., 2010; Naigles, 1990), or conversely, they can map thematic roles in their conceptual space—like agents, patients, and instruments—onto respective arguments in sentences (Perkins et al., 2024). The richness of learners’ conceptual knowledge in informing them about language is furthermore supported by their early acquisition of abstract linguistic expressions like belief verbs (Dudley, 2017) and belief reports (S. Lewis et al., 2017). While the link between the conceptual world and the linguistic signal is by no means transparent across all these cases, learners show biases in drawing on these links in existing knowledge, whether linguistic or conceptual, to constrain their space of hypotheses and arrive at the target generalizations at a striking speed and against all odds.

Taken together, these findings allude to a possibly deeper developmental continuity: the rapid emergence of linguistic knowledge in young learners can be better understood when situated within the context of their already well-established conceptual foundations. Across various learning scenarios, linguistic and conceptual development appear as closely interconnected processes that communicate, support, and inform one another. This perspective provides a powerful lens through which seemingly discrete aspects of development begin to align. It reminds us that language acquisition is not simply about learning words or rules in a vacuum—it is the ongoing

construction of a bidirectional bridge between what learners already know about the world and the linguistic tools they are rapidly learning to use.

6.3.2 The path to generalization: From input to target

While conceptual understanding provides an essential foundation for language learning, it is not sufficient on its own. For learners to comprehend and use sophisticated linguistic forms, they must be able to generalize from the input they receive—input that is often sparse, noisy, or ambiguous. This dissertation explores how young learners form such generalizations under two challenging conditions: when the target input is sparse and when it is unclear. Across both scenarios, the findings show that learners are active inference makers capable of arriving at abstract linguistic generalizations early in development.

In the case of sparse input, this dissertation examines whether children can comprehend complex, sparse structures before the emergence of full-fledged structural knowledge. These instrumental *with*-questions pose formidable challenges. They involve non-local structural dependencies, and they contain argument gaps that occur in structurally optional and infrequent *with*-phrases.

Despite these challenges, 15-month-olds are strikingly capable of responding to questions containing prepositional object gaps correctly. Specifically, prepositional object gaps following preposition *with* are exceedingly rare in input to 9- to 15-month-olds. When such forms do occur, they occur in simpler cases like (47), where *with* follows the verb directly. In contrast, the instrumental *with*-questions that we used during the test, such as (48), featured more complex forms, where a verbal complement intervenes between the verb and the preposition. These more complex structures are virtually non-existent in their early input.

(47) What can Morgan play with?

(48) What did she break the stick with?

Yet remarkably, the absence of such examples in their input does not hinder learners' comprehension. This implies that early learners are not constrained by patterns of string continuations in their input. Instead, they appear to be in a good position to infer what a question is seeking, even if its structure is unfamiliar. Importantly, this pragmatic sophistication at 15 months does not come at the cost of grammatical development. Shortly after, by around 18 months, they start to show clear sensitivity to non-local dependencies (Perkins and Lidz, 2021). This developmental trajectory suggests that in the interim, infants are not just reacting to questions in real time—they are generalizing from simpler, familiar local patterns in the input and building adult-like representations of non-local dependencies.

In the second learning condition with unclear input, the dissertation suggests that learners can acquire the familiarity presupposition of *the*, despite variable conditioning environments of *the* in input. In English, definite determiner *the* implies that the referent is known to the listener, either through prior mention (anaphoric use) or through shared knowledge (hearer-old use). While anaphoric uses of *the* (*a ball...the ball*) may help to scaffold its familiarity presupposition, toddlers' input often features a mix of different uses by their caregivers, sometimes even with more situational uses. If learners primarily attend to discourse properties of referents and ignore their pragmatic contributions to the identification of intended referents, then upon observing highly frequent hearer-old uses of *the* for first mentioned referents, they may infer an incorrect meaning for which *the*-expressions are used for specific referents in the speaker's mind, regardless of whether the referent is mutually known.

Overgeneralization of *the* is, however, not observed in our longitudinal data. By

20 months, learners appropriately restrict their uses of *the*, demonstrating sensitivity to both anaphoric and situational contexts. While *the*-expressions are equally common for first- and subsequent-mention referents in input to 14 months, none of the learners mimic the distribution around 20 months in earliest production. These learning outcomes suggest that learners are not simply tracking surface-level distributional patterns in the input. Instead, they are actively recruiting situational cues to build associations between the distribution of determiner forms and their subtle pragmatic contributions to conversational dynamics, so as to discern the implicit rule governing the different usages of *the*-expressions in communication. The involvement of pragmatic competence in inferring the presuppositions of linguistic forms is indeed consistent with prior findings evincing the early emergence of Theory of Mind (Buttelmann et al., 2009; Onishi and Baillargeon, 2005). Learners' early-developing Theory of Mind may be key to enabling them to infer speakers' intentions and interpret linguistic forms in light of shared knowledge and communicative goals.

Together, findings from these two case studies demonstrate young learners' remarkable capacity to extrapolate beyond existing linguistic data to form abstract, adult-like generalizations. When faced with sparse structural data or unclear referential cues, learners do not content themselves with rote-learned structures; rather, they coordinate information across multiple domains—syntactic, pragmatic, and conceptual—so as to identify the underlying principle. This ability to integrate across domains is not only a hallmark of early language learning but also a window into the powerful inferential systems that support it.

6.3.3 Approaches: Probing competence through performance

Finally, this dissertation sheds light on the methodological challenges and opportunities in assessing young learner’s linguistic competence through experimental tasks. A central insight is that children’s knowledge is often more advanced than their behavior suggests—but accessing that knowledge requires carefully designed methods that account for developmental and contextual sensitivities.

First, the expression of linguistic knowledge in children is fragile and highly context-dependent. Even when children possess the underlying competence, their ability to demonstrate it can be hindered, unless the experimental context aligns with how language is naturally used. This is especially true in the domain of pragmatics, where socially grounded tasks are essential.

Indeed, as communicative acts play a crucial role in helping learners maintain object representations in discourse (Brody and Csibra, 2025), it is equally important to probe their knowledge of discourse referents and expressions for referring back to them in communicative contexts. For example, when investigating referring expressions, tasks must avoid abstract or decontextualized prompts and instead, incorporate real-world social interactions for child participants to be cooperative at the least and for researchers to have even the slightest chance of capturing authentic language use (Moyer et al., 2015; Grigoroglou and Papafragou, 2019; Hamburger and Crain, 1982; Crain et al., 1990).

Second, understanding the developmental trajectory of relevant inferential abilities is critical, both for recalibrating results from prior research and for upgrading future experimental design. Knowing when specific inference skills emerge allows researchers to create age-appropriate tasks that minimize extralinguistic burdens—such as working memory or attentional demands—while still probing the relevant linguistic

knowledge.

The importance of knowing related inferential abilities in predicting the emergence of linguistic competence is reflected in both of our case studies. In comprehending instrumental *wh*-questions with a stranded *with*, learners might find the stranded *with* unexpected if they are making lexically based predications incrementally. That is, as none of the verbs we chose had a tendency to co-occur with a *with*-phrase in declaratives for infant-directed speech, it would be devastating for a young parser to revise the already built structure if it makes strong lexically based structural commitments early on in the sentence (Trueswell et al., 1999; Atkinson et al., 2018). But knowing that 16-month-olds most likely still treat incoming structures veridically (Lidz et al., 2017), we would have hope that learners around this age may still be able to interpret stranded *with*-questions successfully.

Likewise, when understanding the development of reference, we cannot ignore learners' social development alongside their language. Being able to correctly interpret or produce definite expressions depends not only on grasping the linguistic input itself but also heavily on navigating subtle social and pragmatic reasoning. While the task of coordinating non-linguistic information and linguistic expression is by no means easy for learners, we should not ascribe their failure in specific contexts to systematic immaturity without solid evidence. Rather, we should be more skeptical of the deficient-knowledge view when learners clearly show relevant pragmatic competence from early on (Choi et al., 2018; Moll et al., 2008; Saylor and Ganea, 2007) but somehow trip over tasks with extraneous demands.

Taken together, these insights underscore that experimental methods must do more than testing linguistic forms—they must create the right conditions for learners' knowledge to surface. By aligning task design with both developmental readiness and naturalistic use, researchers can more accurately probe the rich, often hidden,

competencies that underlie early language acquisition.

6.4 Future directions

This dissertation explored two key learning problems in early language development: how young learners comprehend structurally complex questions without full-fledged grammatical knowledge, and how they come to identify presuppositional meanings linked to definite *the*-expressions. The findings reveal young learners' early sensitivity to both syntactic structures—such as argument gaps in *wh*-questions—and pragmatic factors, such as the contextual status of referents when producing determiners. Building on these results, I lay out three directions for future research.

6.4.1 Developmental trajectory of question processing

One avenue for further investigation is to trace how children's strategies for processing *wh*-questions evolve over time. While this dissertation shows that 15-month-olds can resolve *with*-questions by relying on a transitivity-based, gap-driven strategy, it remains unclear whether this approach extends to other question types—especially adjunct *wh*-questions like *how*-questions.

A useful next step would be to include *how*-questions as a separate condition using the current experimental design, as shown in (50).

(49) What did she break the stick with?

(50) How did she break the stick?

Unlike argument *wh*-questions (49), adjunct questions like (50) do not involve a missing syntactic argument. If 15-month-olds rely exclusively on detecting argument

gaps to interpret questions, they may struggle to respond to (50). However, if they can successfully interpret *how*-questions, this would suggest that their comprehension strategy is not limited to gap detection. It would also imply that they can process and retain fronted *wh*-phrase when no clear gap is present, at least in the case of adjunct questions. Success with *how*-questions would also be interesting as a possible result, as it would suggest that toddlers hearing argument *what*-questions may be addressing the clause-final gap, knowing that this is sufficient for question understanding.

To probe this further, we can also examine more formal questions like (51).

(51) With what did she break the stick?

This structure—featuring a fronted *wh*-phrase (*with what*)—is rare in everyday speech but can allow us to test whether toddlers have sufficient working memory to maintain the clause-initial *wh*-phrase and still map it onto the corresponding syntactic gap later in the sentence. Success with such items would suggest a more robust capacity for clause-level parsing and working memory than previously assumed at this age.

An equally intriguing question is how early learners handle clause-medial gaps.² This dissertation focuses on instrumental *wh*-questions with a stranded *with*, where the gap is clause-final, similar to direct object questions previously tested in the literature (Gagliardi et al., 2016; Perkins and Lidz, 2020). Learners’ repeated success with clause-final gaps alludes to one possibility—15-month-olds may rely on a naïve, final-gap-driven strategy. This is likely, if an inaudible empty category is psychologically registered in the same way as how clause-final audible elements are. Assuming this, empty categories may also benefit from an *edge advantage* (Newport, 1990).

²This idea was developed by Robert Hopkins, an undergraduate student who took a seminar I taught based on Chapter 2 of the dissertation. Here, I discuss some interesting puzzles based on our discussions.

However, while clause-final grammatical information has been shown to have an edge over medial information (Sundara, 2018), there is evidence suggesting that clause-medial information may also be retained well for structural reasons (Ying et al., 2022). If early learners are indeed guided primarily by final-gap cues, then sentences like (52) may be more difficult for them, if medial gaps are less well-registered given their memory limitations.

(52) Which toy did Kermit pick ___ over there?

Indirect evidence from prior research suggests that 19-month-olds are capable of handling medial gaps (Hirzel, 2022). In an earlier study (Lidz et al., 2017), they often misparsed sentences like (53), interpreting *the tig* as the direct object of *pull* and ignoring *with*, likely due to the strong transitivity bias of *pull*. However, the same learners did not repeat this misparse with direct object questions (54). Instead, they seem to posit a medial gap after *pull*, indicating a developing ability to represent medial gaps.

(53) She is pulling with the tig.

(54) What is she pulling ___ with the tig?

If 19-month-olds are the mature learners who can handle clause-medial structures, an important developmental question is whether 15-month-olds are still operating under the naïve, final-gap strategy. Conversely, if 15-month-olds can resolve medial gaps, as in (52), this would provide compelling evidence for early sensitivity to structural information, even in the face of presumed processing limitations of maintaining sentence-medial information (cf. Sundara, 2018).

Together, these variations would help clarify whether toddlers' comprehension of

wh-questions relies narrowly on syntactic gap-filling, as has been reported so far in the literature (Gagliardi et al., 2016; Perkins and Lidz, 2020). This line of inquiry will offer a promising window into the early architecture of sentence processing and its development over time.

6.4.2 Probing knowledge of discourse referents during childhood

As prefaced in chapter 5, existing studies have yet to clarify when anaphoric uses become cues young learners can recruit for understanding *the*-expressions. How young children represent discourse referents remains an open but important question.

One related area of investigation concerns children’s sensitivity to definite descriptions that are conditioned either by anaphoricity (referring back to something previously mentioned) vs. by situational identifiability. As introduced in chapter 5, in German, the dative singular definite determiner *dem* can take different forms when following a preposition *zu* ‘to’, depending on whether the referent has a prior discourse antecedent (*zu dem* ‘to the_{strong}’) or is merely familiar from the current situation (*zum* ‘to-the_{weak}’). If children are more sensitive to anaphoric, strong definite forms—with their surface signals of antecedents—we would expect them to comprehend strong definites in German earlier than weak definites (the contracted form) that hinge on sensitivity to situational knowledge.

Another promising direction of research is to directly investigate how infants and toddlers retrieve discourse referents. One study by Saylor and Ganea (2007) suggests that by 12 months, infants seem to rely heavily on spatial information to identify previously mentioned referents intended by specific speakers. When these referents are spatially displaced, infants struggle to retrieve them. This raises the question of whether discourse information alone can support young learners’ referent retrieval

when the referent is no longer visible. To address this, neural markers of reference resolution offer a promising approach. For instance, Brodbeck, Gwilliams, and Pytkäinen (2015) found that reference resolution in adults is associated with a negative evoked potential at posterior electrodes contralateral to the spatial location of a previously seen referent. This method could provide an ideal way to study resolution of referents that have previously been seen but are visually absent at the moment. Crucially, participants' reference resolution can be measured without requiring explicit behavioral responses. However, one drawback is that this technique might work less well with older toddlers, who tend to become fussy and less cooperative during longer experiments.

6.4.3 Acquiring referring expressions in determiner-less languages

A third direction for future research concerns how children acquire referential expressions in languages that lack dedicated articles or determiners for definiteness. In English, determiners like *the* and *a* serve as explicit morphological markers for whether a referent is known or new to the listener. However, many of the world's languages—including Mandarin—do not encode definiteness morphologically. For example, Mandarin speakers typically use bare nouns to express definites (Cheng and Sybesma, 1999; Jenks, 2018; Dayal and Jiang, 2022), as in (55). In contrast, indefinites are usually marked by combining numerals with classifiers (56).

- (55) hézi bújiàn le
 box gone PFV
 'The box is gone.'

- (56) yí-gè hézi bújiàn le
 one-CL box gone PFV
 'A/one box is gone.'

Interestingly, Mandarin presents a learning scenario where indefinites are marked whereas definites are unmarked. The absence of explicit definiteness marking raises an important question: how do learners acquiring Mandarin or similar languages come to produce referential expressions when the intended referent is unique vs. non-unique? Do they use bare forms for definite descriptions as the theoretical literature describe, or do they develop an alternative way to be informative when the intended referent is definite?

One possible hypothesis is that young learners may start by using a demonstrative such as *nage* ‘that-CL’ in addition to the noun phrase to mark definiteness in an over-informative manner. In this view, to arrive at the target usage, they will need to identify how bare nouns are the right fit on a scale of referential forms to convey definiteness and delete explicit markings on noun phrases in their adaptation of referring expressions.

Alternatively, young learners may converge on the right bare form for marking definiteness, even earlier than their English-learning peers. Regarding this possibility, the lack of explicit marking may be pragmatically motivated, in the sense that what is presupposed or taken for granted is silent by default. This is possible, considering how Mandarin also allows the omission of arguments, if they are given by prior discourse or the situation (C.-T. J. Huang, 1984; I. Roberts and Holmberg, 2009).

Ultimately, studying how learners manage reference in languages without determiners will provide crucial insights into how learners acquiring different languages identify ways to mark definiteness. This will further clarify the extent to which learners are actively attending to either anaphoric or situational cues to identify how definiteness is expressed, even when it may take a bare form.

6.5 Conclusion

This dissertation explored how young language learners navigate complex linguistic forms, with a focus on two key domains: question comprehension and determiner production. Across these case studies, the evidence shows that even in the face of sparse and unclear input, children actively infer structure and meaning, drawing on linguistic structures, conceptual understanding, and contextual cues to form generalizations that extend beyond their input.

Together, these findings paint a developmental picture in which language acquisition is driven by a flexible, inferential, and resilient cognitive system—one that integrates syntax, conceptual reasoning, pragmatics, and real-world knowledge from an early age.

Beyond the empirical contributions, this work highlights the critical role of methodology in developmental research. Accurately assessing children’s linguistic competence—especially in context-sensitive areas like pragmatics—requires careful experimental design that accounts for performance factors and task demands.

In sum, this dissertation illustrates how language learning unfolds as a dynamic, inferentially rich process. Young learners go beyond surface linguistic forms, engaging with deeper patterns in the language—they interpret, infer, and find ways to connect the dots, even when those dots are few and far between. This capacity to generate structure and meaning from limited input reflects the remarkable ingenuity of the human mind at work from the very beginning.

Appendix A: Data from Study 1

Corpus	Education	DO	PO	S	DO%	PO%	S%
c1	higher	57	17	5	72.2	21.5	6.3
d1	higher	51	8	2	83.6	13.1	3.3
f1	higher	135	7	10	88.8	4.6	6.6
s1	higher	135	10	3	91.2	6.8	2.0
s2	higher	228	5	10	93.8	2.1	4.1
s3	higher	51	5	0	91.1	8.9	0
v1	higher	90	9	2	89.1	8.9	2.0
w1	higher	43	4	29	56.6	5.3	38.2
f2	lower	68	11	1	85.0	13.8	1.3
i1	lower	106	5	3	93.0	4.4	2.6
j1	lower	132	13	4	88.6	8.7	2.7
m2	lower	66	1	0	98.5	1.5	0
q1	lower	24	4	1	82.8	13.8	3.4
t1	lower	55	7	8	78.6	10.0	11.4
v2	lower	137	22	1	85.6	13.8	0.6
w3	lower	40	7	9	71.4	12.5	16.1

Table A.1: Proportion of *what*-questions by Education Level

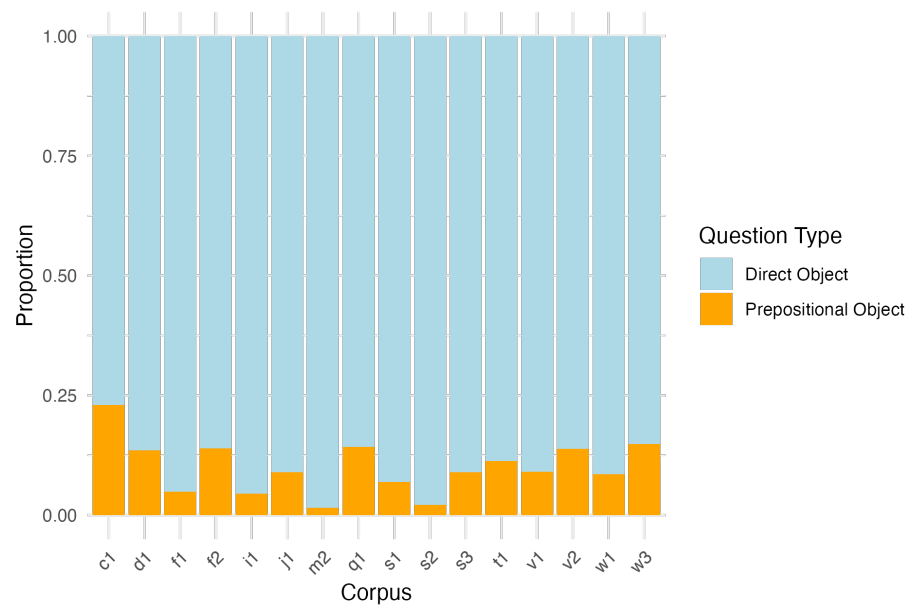


Figure A.1: Proportion of DO vs. PO *what*-questions by Corpus

Child	DO	PO	Total
c1	57	17	74
d1	51	8	59
f1	135	7	142
s1	135	10	145
s2	228	5	233
s3	51	5	56
v1	90	9	99
w1	43	4	47
f2	68	11	79
i1	106	5	111
j1	132	13	145
m2	66	1	67
q1	24	4	28
t1	55	7	62
v2	137	22	159
w3	40	7	47
Overall	1418 (91.3%)	135 (8.7%)	1553

Table A.2: Distribution of DO vs. PO *what*-questions (with subject questions removed)

Corpus	with+DP	with+ $\emptyset$	(with+DP)%
c1	55	5	91.7%
d1	96	7	93.2%
f1	73	13	84.9%
s1	66	10	86.8%
s2	80	0	100%
s3	14	0	100%
v1	137	4	97.2%
w1	77	2	97.5%
f2	53	3	94.6%
il	75	8	90.4%
j1	37	3	92.5%
m2	44	6	88%
q1	22	2	91.7%
t1	74	6	92.5%
v2	70	5	93.3%
w3	99	7	93.4%
Overall	1072	81	93%

Table A.3: Transitive vs. intransitive uses of preposition *with* across corpora

condition	window	mean	sd	stats
Wh	base	0.53	0.16	$t(11) = 0.65, p = .265$
Wh	test1	0.51	0.20	$t(11) = 0.15, p = .441$
Wh	test2	0.62	0.19	$t(11) = 2.10, p = .03$
Con	base	0.47	0.15	$t(11) = -0.70, p = .749$
Con	test1	0.47	0.14	$t(11) = -0.88, p = .802$
Con	test2	0.45	0.12	$t(11) = -1.46, p = .914$

Table A.4: One sample t-tests: Proportion of target looking above chance level

Appendix B: Stimuli for Study 2

This appendix presents stimuli for testing instrument questions.

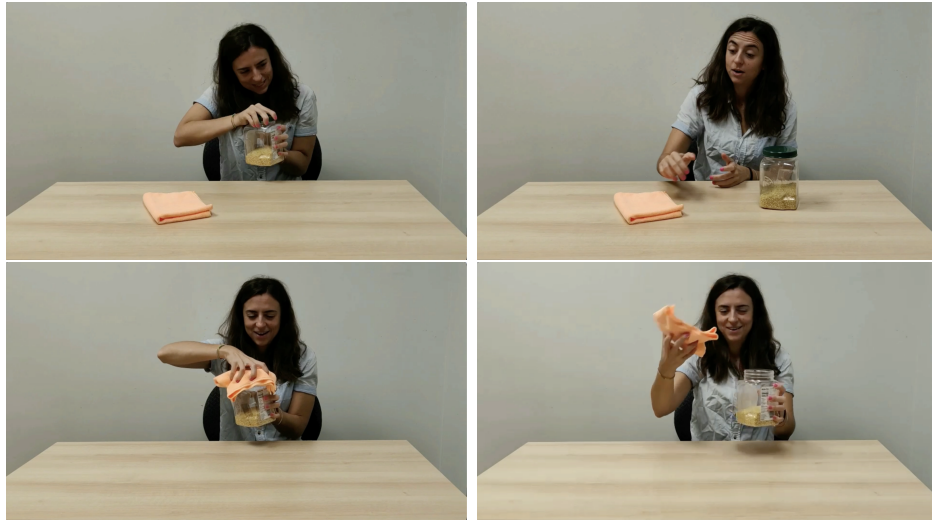


Figure B.1: Images of an experimental video for an OPENING scene

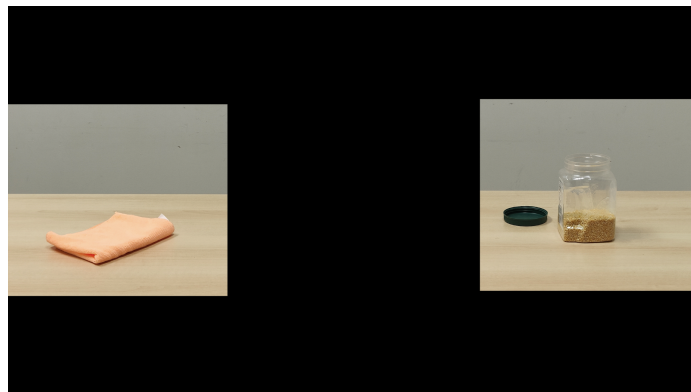


Figure B.2: Image of a test phase for the OPENING scene

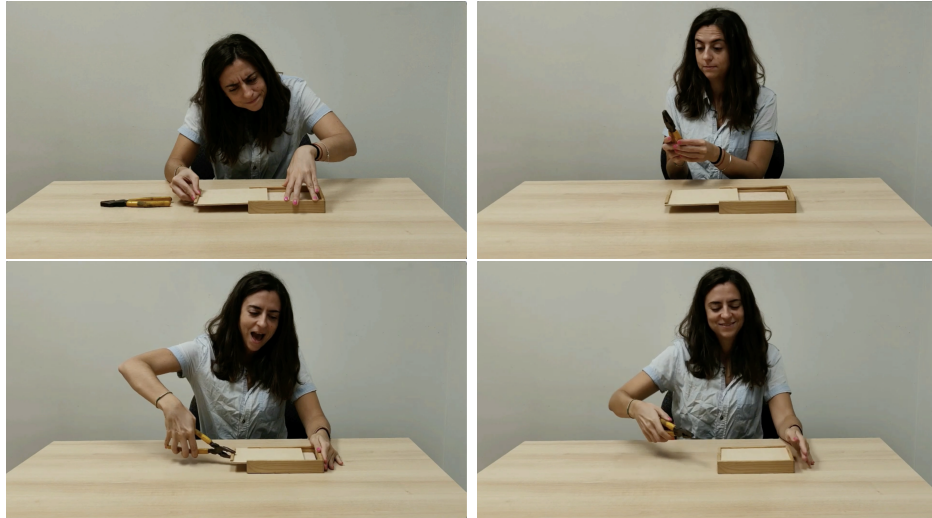


Figure B.3: Images of an experimental video for a CLOSING scene

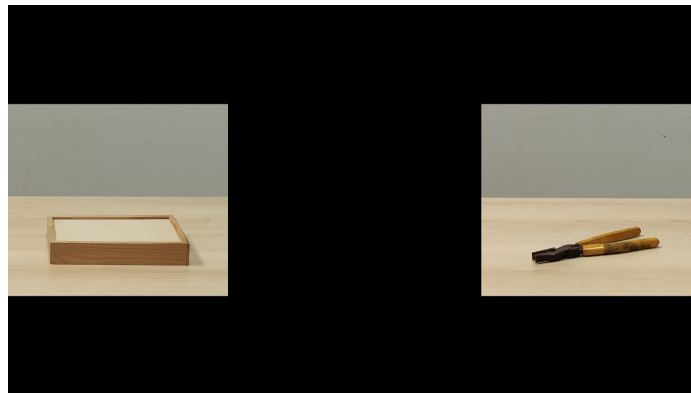


Figure B.4: Image of a test phase for the CLOSING scene



Figure B.5: Images of an experimental video for a BREAKING scene

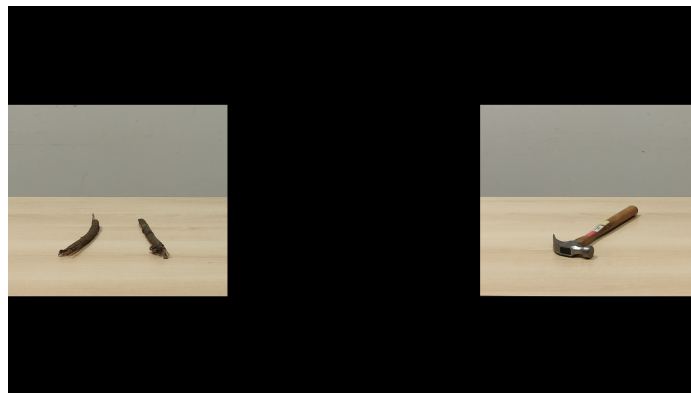


Figure B.6: Image of a test phase for the BREAKING scene



Figure B.7: Images of an experimental video for a CLEANING scene



Figure B.8: Image of a test phase for the CLEANING scene



Figure B.9: Images of an experimental video for a FIXING scene

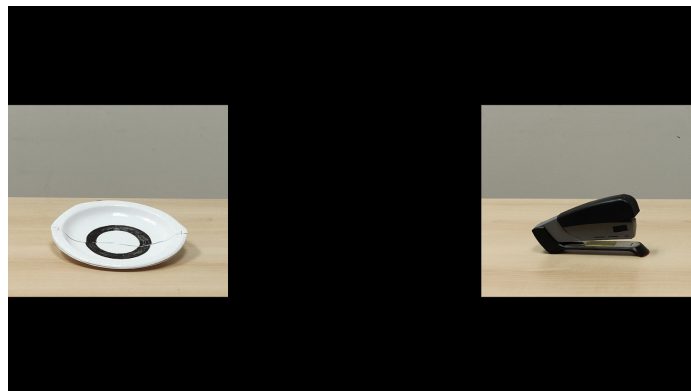


Figure B.10: Image of a test phase for the FIXING scene

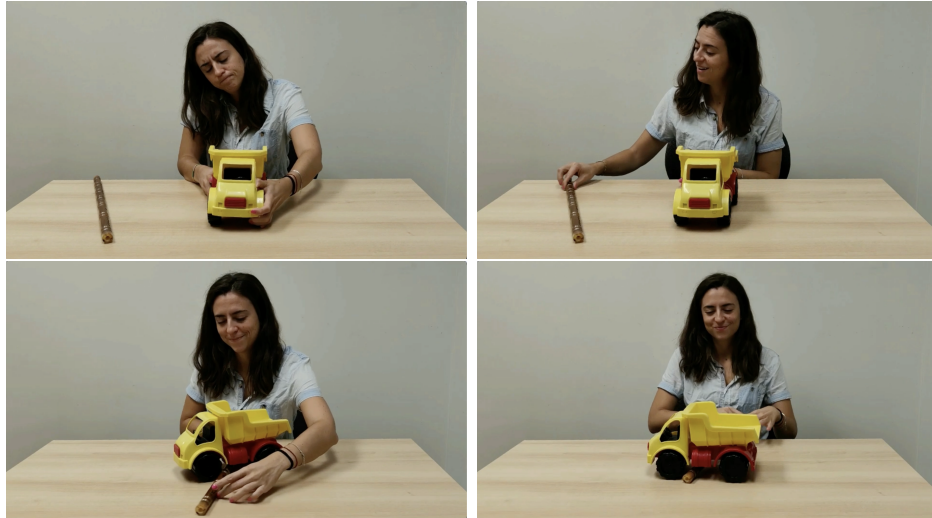


Figure B.11: Images of an experimental video for a TURNING scene



Figure B.12: Image of a test phase for the TURNING scene

Appendix C: *The*-misuse for 2-year-olds' corpora in Study 4

Below are cases of *the*-misuse found in 2-year-olds' corpora for Study 4.

(1)

COLIN: What were you doing (.), Eve?

CHILD: In the basement.

CHILD: Here's window.

CHILD: Have a get the window.

CHILD: Have a get the window.

RICHARD: Which window?

CHILD: This window.

CHILD: Where window.

(Brown/Eve/020000b, 02;00;00)

(2)

CHILD: Men (.) please.

MOTHER: Men (.), please.

MOTHER: Where will I get them?

CHILD: In the ...

CHILD: Man in the men in the box.

MOTHER: In what box?

MOTHER: xxx another piece of cheese (.) to make another man?

MOTHER: Will you eat it?

CHILD: And I go eat it too.

(Brown/Eve/020100b, 02;01;00)

(3)

MOTHER: Oh that was good.

CHILD: Yeah.

MOTHER: What room does the dog stay in?

CHILD: Here.

CHILD: No dog stay in the room.

CHILD: Don't talk stay in the room.

MOTHER: Which room?

CHILD: On the chimney.

MOTHER: On the chimney?

(Suppes/020115, 02;01;15)

(4)

CHILD: Yeah.

MOTHER: That was fun wasn't it.

MOTHER: Tell Linda where you went.

MOTHER: Where did you go?

CHILD: In the airplane.

MOTHER: You went in an airplane?

LINDA: Where else?

CHILD: Went in the train.

MOTHER: In a train too.

(Suppes/020406, 02;04;06)

(5)

MOTHER: You wanna come on the what?

CHILD: Platforms.

MOTHER: On the platform?

CHILD: Yes.

CHILD: I wanna go on the platform.

MOTHER: What platform?

CHILD: The platform at the circus.

CHILD: And um...

MOTHER: Oh, you mean where the elephant was?

(Valian/14b, 02;04;26)

(6)

MOTHER: Okay.

MOTHER: I want you to try to be a little bit more with what you're making, okay?

MOTHER: You can play, but you don't hafta throw the things all over.

MOTHER: That's not necessary.

CHILD: I wanna see the teacher again, Mommy.

MOTHER: Which teacher?

CHILD: That teacher.

MOTHER: Observer?

MOTHER: Oh, you mean Susan?

(Valian/9a, 02;07;25)

(7)

MOTHER: Just a moment.

CHILD: What's the people doing.

MOTHER: What people?

CHILD: The people there.

MOTHER: There aren't any people out there.

MOTHER: I was peeping out at Sarah.

CHILD: Hmm?

(Brown/Eve/020200b, 02;02;00)

(8)

MOTHER: What color is that dump truck?

CHILD: Yellow.

MOTHER: Nope.

CHILD: Orange.

MOTHER: Right.

CHILD: But where the blue one?

MOTHER: What?

CHILD: Where's the blue one?

MOTHER: Which blue one?

CHILD: The other one.

CHILD: Hafta go turn the other light on.

(Valian/13a, 02;04;02)

(9)

INVESTIGATOR: I think the porcupine is behind you.

INVESTIGATOR: Right back there.

CHILD: Where's the truck?

CHILD: Where's the truck?

CHILD: Where's the truck?

CHILD: Hey, where's the truck?

CHILD: I need xxx.

INVESTIGATOR: Where's the what?

MOTHER: The truck?

CHILD: Yeah.

MOTHER: There's two trucks.

(Valian/10b, 02;05;07)

(10)

MOTHER: What did you find for my lunch?

CHILD: Corn.

MOTHER: Oh.

MOTHER: We need butter then.

MOTHER: Find me some butter to put on it.

CHILD: Where's the corn?

MOTHER: The corn on the cob?

MOTHER: Maybe it's in the box.

MOTHER: See if you can find it.

MOTHER: Do you have more food in here?

MOTHER: Oh (.) there's a lot of food in there.

(Suppes/021112, 02;11;12)

Appendix D: Stimuli for Study 5

This appendix includes instructions to participants and test stimuli for both definite and indefinite uses.

Participants first read the following instructions before getting to the practice trials.

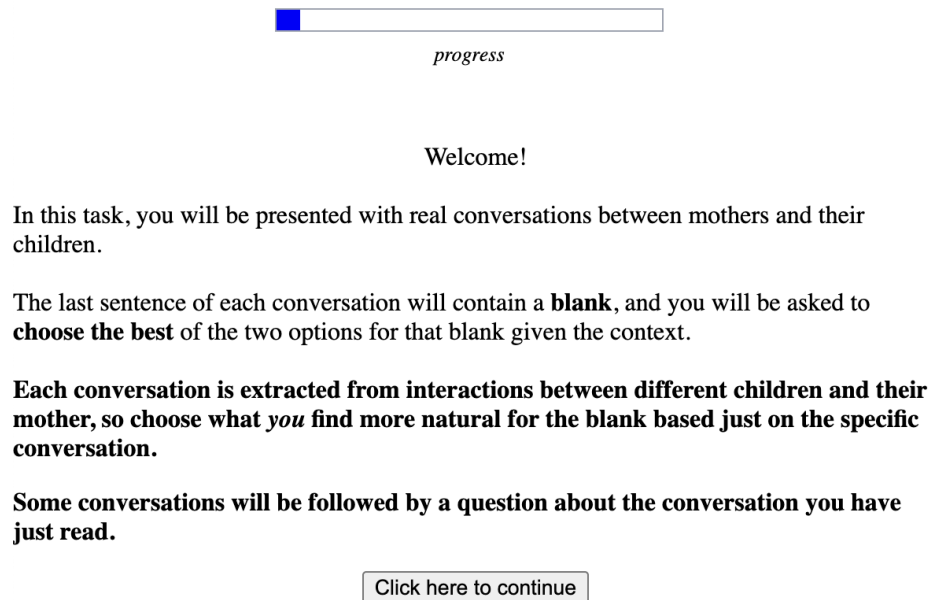


Figure D.1: Instructions to participants before the experiment

Participants are then randomly assigned to one of the six conditions. The conditions adopt a 2 (speaker: child vs. mother) x 3 (child age in corpora: 2- vs. 3- vs. 4-year-olds) design.

D.1 Child determiner uses in 2-year-old corpora

age	speaker	item	ans	context
2	chi	1	the	CHILD: Hey. MOTHER: You want me to tell you? MOTHER: Or are you gonna tell me what's going on? CHILD: Hey. MOTHER: Hey, what's this? MOTHER: Let's start at the beginning. MOTHER: Ya know, this isn't the beginning. MOTHER: No. MOTHER: Let's find the beginning. CHILD: This is __ beginning.
2	chi	2	the	MOTHER: Oh, wait a minute. MOTHER: Where's that kitty? CHILD: This one ... MOTHER: Remember who this is? CHILD: Yeah. MOTHER: Who's that? MOTHER: The mailman? CHILD: A policeman. MOTHER: Right. CHILD: And that ...and that's __ kitty.

2	chi	3	the	MOTHER: Or are you just gonna take it apart? CHILD: I'll put this right here. MOTHER: Mhm. MOTHER: You know what else? CHILD: Yeah. MOTHER: I'll bet there's something in here. MOTHER: That's the engine. MOTHER: That's what makes the car work. CHILD: I'll close this. CHILD: I'll just close __ trunk.
2	chi	4	the	CHILD: This'll be a plate. CHILD: This my plate. MOTHER: Mhm . CHILD: And these are plates. MOTHER: Mhm . CHILD: Yeah . CHILD: And these are big plates. CHILD: Put chairs. CHILD: Put the plates over here. Put these over here. The fork goes on the left side. CHILD: And __ spoon goes on this side.

2	chi	5	the	CHILD: Look at these money. MOTHER: Okay, there's a penny on the floor behind you. MOTHER: Do you see it? CHILD: Where? MOTHER: Here. CHILD: What is this? CHILD: What's this in here? MOTHER: Oh, well, on a real cash register the tape would come out there. MOTHER: (Re)member the white piece of paper? CHILD: Just close ___ door.
2	chi	6	the	CHILD: What is it? MOTHER: Cereal? CHILD: What's inside? MOTHER: Nothing really. CHILD: I put tea inside. MOTHER: Tea? CHILD: No. CHILD: This is milk. MOTHER: Okay. CHILD: Take ___ top off.

2	chi	7	the	CHILD: Who's in carriage? MOTHER: Who do you think that is? CHILD: Kitty cat. MOTHER: Right. MOTHER: Have you ever seen a little kitty cat outside? CHILD: And there's the kitty cat there. MOTHER: Right, this book says The great cat chase. MOTHER: So here's the little girl who's playing mommy, and she's picking the flowers. MOTHER: Look, what's the cat doing? CHILD: Hiding and seeking on __ tree.
2	chi	8	the	CHILD: Six. CHILD: Seven... MOTHER: Yeah. CHILD: Seven. MOTHER: Are you having fun with that? CHILD: And I close this trunk. MOTHER: Okay. CHILD: And this will go in the garage. MOTHER: Okay. CHILD: Uh someone drive __ car.

2	chi	9	the	MOTHER: Alright, I'll finish it. MOTHER: And then you'll put the tools away? MOTHER: Okay. CHILD: I wanna do this. MOTHER: Put the tool ... put the tools in. MOTHER: And then we'll put this away. CHILD: And then I do want. MOTHER: And now... MOTHER: Okay, then we'll look and see what's there okay? CHILD: That's ___ garage.
2	chi	10	the	CHILD: Uh someone drive the car. CHILD: And everybody going to the garage to here. MOTHER: Okay. CHILD: Goodbye garage. CHILD: Goodbye. CHILD: Goodbye. CHILD: Goodbye. CHILD: And this'll be the garage. CHILD: That will be all three. CHILD: This'll be ___ garage.

2	chi	11	the	MOTHER: Look! MOTHER: Right back into the tunnel there goes the Googey. MOTHER: In goes the lady after him. MOTHER: After the Googey. MOTHER: And in goes the policeman after the lady. CHILD: Where's the Googey? MOTHER: And here's the Googey. MOTHER: Who caught Googey? MOTHER: The CHILD: He's carrying ___ boy.
2	chi	12	the	CHILD: Is this the table? MOTHER: Mhm. CHILD: I do these here? MOTHER: Okay. CHILD: Are these doors? MOTHER: Mhm. CHILD: Are these closed? MOTHER: Mhm. CHILD: Tight? CHILD: Is this ___ couch?

2	chi	13	the	CHILD: Oh ... CHILD: And this is the bed for ... for the child that goes in there. MOTHER: Could this be the children? CHILD: Where? MOTHER: In the back? CHILD: Yeah. CHILD: Where ... where's the papa? MOTHER: Well , this could be the papa. CHILD: Yeah. CHILD: That's __ papa.
2	chi	14	the	MOTHER: So here's the ...the little girl who's playing mommy, and she's picking the flowers. MOTHER: Look , what's the cat doing? CHILD: Hiding and seeking on the tree. MOTHER: He's hide-and-seeking under the tree? CHILD: Yeah. MOTHER: That seems to be what he's doing. CHILD: Oh ... oh. MOTHER: That's a good word, hide and seeking. CHILD: Oh there's the doggy. CHILD: There's __ kitty cat.

2	chi	15	the	CHILD: There's kitty cat. MOTHER: It sure is. MOTHER: Ya wanna see what happens with this kitty cat? CHILD: Yeah. MOTHER: What's ... MOTHER: Aah oo... he's so upset because she doesn't know where her baby kitty cat is. MOTHER: Look, who's that? MOTHER: Who's that there? CHILD: Who's that there? CHILD: ___ lady.
2	chi	16	the	MOTHER: And turn, see? CHILD: Okay , I'm ... MOTHER: I'm turning it the wrong way. MOTHER: I forgot .. MOTHER: That you're a leftie. MOTHER: Here, this way, okay? CHILD: I have people. MOTHER: You wanna put the people in? CHILD: Yeah. CHILD: They hafta get to ___ garage.

2	chi	17	the	MOTHER: You can turn the can on the side, if you wanted. MOTHER: If it's easier for you can move it. MOTHER: That's it. MOTHER: Turn it ... CHILD: I can't. MOTHER: Turn it, all the way around ... MOTHER: Yes you can. MOTHER: Here. MOTHER: Turn it this way. CHILD: I'll open __ trunk.
2	chi	18	the	CHILD: Sit on the chair. MOTHER: Let me ... MOTHER: I.. I don't like the way you say it. CHILD: Sit on the chair please. CHILD: Let's look something. MOTHER: Okay. MOTHER: I'm sitting up. MOTHER: You wanna read me the story? MOTHER: Or you wanna read the story? CHILD: You read __ story.

2	chi	19	the	CHILD: Is this the table? MOTHER: Mhm. CHILD: I do these here? MOTHER: Okay. CHILD: Are these doors? MOTHER: Mhm. CHILD: Are these closed? MOTHER: Mhm. CHILD: Tight? CHILD: Is this ___ couch?
2	chi	20	the	MOTHER: Do you know what this is? MOTHER: This is play money. CHILD: Oh. MOTHER: Okay? MOTHER: You wanna stay by the cash register. MOTHER: And I'll give you the money for the things that we bought? MOTHER: Okay? CHILD: Here. MOTHER: What? CHILD: In ___ garage.

2	chi	21	a	MOTHER: Do you know that this car comes apart? MOTHER: Do you know that you can unscrew this right here? CHILD: Yeah. MOTHER: That's right yeah look at that. MOTHER: Some more. MOTHER: Now you hafta keep turning in the same direction. CHILD: And this. MOTHER: Yep. CHILD: And then ... CHILD: I . . . I have ___ screwdriver.
2	chi	22	a	MOTHER: Where do they sit? CHILD: These go. CHILD: These are ... driving. MOTHER: He's driving? CHILD: Yeah. MOTHER: Who's driving? CHILD: This man is drive. MOTHER: This man is driving? MOTHER: I think that's a lady. CHILD: This ___ boy?

2	chi	23	a	MOTHER: Is that where the driver sits? MOTHER: Is it okay if I read this book? MOTHER: While you're doing that? CHILD: Yeah. CHILD: Where does this go? MOTHER: Oh, you wanna put it together? MOTHER: Uum. CHILD: Yeah. MOTHER: Is that a workbench? CHILD: Yeah, __ workbench.
2	chi	24	a	MOTHER: Will you give it to me for free? MOTHER: No? MOTHER: What kind of money is that? MOTHER: Do you know what that's called? CHILD: Cash register. MOTHER: Yeah well, this is a cash register. MOTHER: Where is the money? MOTHER: Show me the money. MOTHER: And what's this called? CHILD: __ dime.

2	chi	25	a	MOTHER: What is he doing? CHILD: What's Puddy doing? MOTHER: There's Puddy right there. MOTHER: But look at the little boy! MOTHER: The little boy is in the carriage. MOTHER: I don't know. MOTHER: This is a funny mixed up story. MOTHER: Oh, look it, they're pretending now to have a tea party. MOTHER: Do you know what she's wearing? CHILD: ___ hat.
2	chi	26	a	CHILD: Oh. MOTHER: It's a little tiny paper napkin box. MOTHER: Okay? CHILD: Yeah. MOTHER: What.. MOTHER: What else should we buy? CHILD: Fruit. MOTHER: There's fruit in here? CHILD: Get fruit ... CHILD: First get fruit at ___ store.

2	chi	27	a	CHILD: I can't. MOTHER: Okay , so leave it the way it is, okay? CHILD: We'll eat later. MOTHER: What? CHILD: We'll eat later. MOTHER: We'll eat later? CHILD: Yeah. MOTHER: Okay, what should we do now? CHILD: This ... CHILD: Huh, __ story.
2	chi	28	a	CHILD: Yeah. CHILD: Here's some dimes. CHILD: Some quarters. MOTHER: What's this? CHILD: It's a quarter. MOTHER: What's this? CHILD: Penny. MOTHER: Penny. MOTHER: And what's this? CHILD: __ dime.

2	chi	29	a	CHILD: You drink it. MOTHER: Can I have tea? CHILD: More tea with the coffee. MOTHER: I really wanted tea. MOTHER: Oh, together? MOTHER: Good. MOTHER: Can you mix it up a little? MOTHER: Stir it with a spoon, maybe? CHILD: oh.. oh , where's the spoon? CHILD: Here's __ knife.
2	chi	30	a	CHILD: Yeah.. MOTHER: Yeah? MOTHER: I'll go get it. MOTHER: Here, let's put the car back where it belongs. MOTHER: Okay. CHILD: No! MOTHER: You don't wanna put the car there? MOTHER: Okay. CHILD: No. CHILD: Wanna play with it and read __ book, okay?

2	chi	31	a	MOTHER: Where is the money? MOTHER: Show me the money. MOTHER: And what's this called? CHILD: A dime. MOTHER: No. MOTHER: That's not a dime this ... CHILD: Is penny. MOTHER: A penny. MOTHER: This is a dime. CHILD: __ dime?
2	chi	32	a	MOTHER: You're still paying? MOTHER: Let me see if we can open up the drawer in that cash register, and you can put the money in there. MOTHER: I think that you can open up the drawer somehow. MOTHER: Oh, here it is. MOTHER: Look at that! MOTHER: Look what's in there. MOTHER: What's this? MOTHER: Look. MOTHER: What's that? CHILD: __ quarter.

2	chi	33	a	MOTHER: I don't think we're sposta touch that, sweet-heart. MOTHER: That belongs to... the people that... who own this playroom. MOTHER: Here's a kleen(ex)... a hankie. MOTHER: Here you go. MOTHER: You wanna wipe it? MOTHER: That's it. MOTHER: Want me to hold it for you after? CHILD: No. MOTHER: Okay, just leave it there? CHILD: Look, that's ___ nurse.
2	chi	34	a	CHILD: Oh. MOTHER: To the middle. CHILD: I'll put in... in the middle. MOTHER: You'll put in the middle? CHILD: And I'll just.. Put'em ... CHILD: And just I'll put this around, around. MOTHER: Be careful. CHILD: I'll see. CHILD: Okay, knife. CHILD: This is ___ knife.

2	chi	35	a	MOTHER: Oh, I'm sorry. CHILD: I wanna put that man in there. CHILD: That man in there. CHILD: That's right. CHILD: And this... CHILD: That's right. CHILD: This one sits in the back and this one sits in the front and drives. MOTHER: Mhm. MOTHER: Where are they going? CHILD: For __ ride.
2	chi	36	a	MOTHER: Okay. CHILD: There it is. CHILD: ... Do... MOTHER: Is it is it too tight? MOTHER: You want me to start it? CHILD: Start with this. MOTHER: Okay? CHILD: Is that in there? MOTHER: That's right. CHILD: I have __ screwdriver.

2	chi	37	an	MOTHER: You're gonna spill the milk? MOTHER: I think you're gonna make a mess. CHILD: No. MOTHER: Can you count these bottles? CHILD: One, two. MOTHER: One, two ... CHILD: One, two, three, four, five. MOTHER: Oh, I think you have an extra one. MOTHER: Count again. CHILD: And I have __ extra one.
2	chi	38	a	MOTHER: He's wearing a hat. MOTHER: He looks like somebody in one of the books that we have at home. MOTHER: Peter the Policeman. MOTHER: Remember? CHILD: Peter Policeman. MOTHER: Well, we can call him Happy the Policeman what is he blowing? CHILD: A whistle. MOTHER: Yeah, and he's stopping the traffic here. MOTHER: And who comes along? CHILD: __ boy.

2	chi	39	a	MOTHER: Here. MOTHER: You hold this here. MOTHER: And you unscrew the wheel. MOTHER: Oo, it's on very tight. MOTHER: You wanna do it? CHILD: Yeah. MOTHER: Turn it. CHILD: I'll do it. MOTHER: See if you can get the tire off. CHILD: I'll close ___ trunk.
2	chi	40	a	CHILD: I wanna do this there. MOTHER: What? CHILD: A knife. MOTHER: Mhm. CHILD: A knife and a spoon. MOTHER: We gonna set the table? CHILD: And a fork. MOTHER: Mhm. CHILD: And a spoon. CHILD: And here and ___ fork over here.

D.2 Child determiner uses in 3-year-old corpora

age	speaker	item	ans	context
-----	---------	------	-----	---------

3	chi	1	the	MOTHER: We could pretend that this was a store, right? CHILD: Yeah. MOTHER: And are you gonna pay me for the cookies? CHILD: Yup. MOTHER: Oh okay. MOTHER: Okay... MOTHER: But then if you're gonna to pay me for the cookies, then i hafta work the cash register, right? CHILD: No. MOTHER: I'll hafta put the money in the cash register when you pay me. CHILD: I'm... no, I'm ___ lady.
---	-----	---	-----	--

3	chi	2	the	MOTHER: You know what? CHILD: What? MOTHER: She's strolling her cat in a baby carriage, isn't she? CHILD: Yeah. MOTHER: That's pretty silly, huh? CHILD: Yes. CHILD: Why? MOTHER: Okay, well Charlie, she stops to pick a flower. MOTHER: And what happens? CHILD: ___ baby gets out.
3	chi	3	the	MOTHER: These things come out here like that. CHILD: Yeah but what about these? MOTHER: Oh. MOTHER: Can you turn ... MOTHER: I think the wheels will come off if you turn the screws on the wheels too. MOTHER: Can you do that? MOTHER: Oop! MOTHER: Get down. CHILD: How can we put it together again? CHILD: Put ___ wheel back on.

3	chi	4	the	MOTHER: Huh? CHILD: Was he scared? MOTHER: I don't know. MOTHER: You know what ? MOTHER: Let's go on the floor cause that book keeps sliding. MOTHER: I wanna see what happens at the end. MOTHER: Come on. MOTHER: Come sit down here with mommy. CHILD: No. CHILD: I wanna sit up on ___ chair and open my store.
3	chi	5	the	MOTHER: The one that has the spider in it ... all folded up like this. CHILD: Oh. MOTHER: So I won't give you one that's all folded up. CHILD: Why? MOTHER: Because that one has the spider in it. MOTHER: Why are you holding that one all up? CHILD: Cause I don't wanna have a spider in it. MOTHER: Yeah but don't you need it for your nose? CHILD: No I don't. CHILD: I put ___ booger in it.

3	chi	6	the	MOTHER: That's right. MOTHER: Very good. MOTHER: You're almost through taking it apart. MOTHER: Just a couple more things. MOTHER: Mhm. MOTHER: Okay now you can do it with your hands a little. MOTHER: That's right. MOTHER: Yep. MOTHER: You've got it. CHILD: I'm gonna give you all your money at ___ store.
3	chi	7	the	MOTHER: Yeah. CHILD: Hey this is a great toy. MOTHER: That's a really good toy, isn't it? CHILD: Right. CHILD: You pretend you're walking along over there. MOTHER: You want me to walk along over there? CHILD: Yeah. MOTHER: Okay. MOTHER: And then what happens? CHILD: Then you come over here to ___ store.

3	chi	8	the	MOTHER: Now ...um... okay so you wanna put a wheel back on? MOTHER: Is that what you wanna do? CHILD: Oh! CHILD: Wait a minute. CHILD: These things. MOTHER: What's that? MOTHER: Goodness! MOTHER: I wonder where that came from, chuck. MOTHER: Something I don't understand. CHILD: Where ___ man sit?
3	chi	9	the	CHILD: Yes. CHILD: Where'd this go which ... where'd... MOTHER: But you know what you hafta check inside first? MOTHER: Why don't you open up and see if the bolt is there? MOTHER: That little yellow piece. MOTHER: Oopsie. MOTHER: We're missing the little bolt. CHILD: Which one? MOTHER: How about if we put that one in that square little hole there? CHILD: ___ rectangle.

3	chi	10	the	MOTHER: How come mine won't go in? CHILD: Ah. CHILD: I think we hafta put these tools back. MOTHER: That is that? CHILD: Tools. MOTHER: Hm? CHILD: I think we hafta take all the... MOTHER: These are okay. CHILD: I hafta put the wheel back in. CHILD: So first we hafta put in ___ screw and do it.
3	chi	11	the	MOTHER: Look at the cowboy boots. MOTHER: And they're walking around looking for the kitty. MOTHER: And the kitty's gonna go hide inside that tire. MOTHER: In that pipe. MOTHER: And look. MOTHER: What are they doing? MOTHER: Are they crawling inside? CHILD: Oh. MOTHER: Yeah. CHILD: The last is ___ cat.

3	chi	12	the	CHILD: This is Melanie. MOTHER: Hah. CHILD: That's powder for you? MOTHER: That's to wash the baby's clothes with, see? MOTHER: Ivory snow. MOTHER: Soft, to make baby's clothing very soft. CHILD: And then they'll be snuggly afternoon. MOTHER: That's right. MOTHER: Yeah. CHILD: And then when the afternoon comes, it will probably be... be dry, and ___ baby could wear them.
3	chi	13	the	MOTHER: Got it! MOTHER: Wait. MOTHER: Mommy was wrong. CHILD: No... Now... Now... can we put this back in? CHILD: The car is parking? MOTHER: In the back here? MOTHER: Yeah. MOTHER: See if you can do that. CHILD: Okay. CHILD: I'll put ___ tire back in.

3	chi	14	the	MOTHER: But tell me about it. MOTHER: You're gonna tell me a story. CHILD: Oh. MOTHER: Tell me what she's doing. MOTHER: Is she taking the kitty cat for a walk in...in the... in the carriage? CHILD: Yeah. CHILD: And ... CHILD: And the ... MOTHER: And ... CHILD: And ___ kitty fell out.
3	chi	15	the	CHILD: Now I'll do the other side. MOTHER: Mhm. CHILD: Oh. CHILD: There are all kinds of things there first I'm gonna do this wheel. CHILD: Whoops! CHILD: There. MOTHER: Uhhuh. CHILD: I can see half of that. MOTHER: You can see some of the store. CHILD: I can... I can see half of ___ ca... cash register.

3	chi	16	the	CHILD: The rectangle. MOTHER: I'd say that's good. MOTHER: Except it's gotta fit in, doesn't it? MOTHER: Does it fit? CHILD: Yep. MOTHER: Good. CHILD: In here? MOTHER: Uhhuh. MOTHER: Now what're we missing? CHILD: ___ screw.
3	chi	17	the	MOTHER: The other way. MOTHER: That's it. MOTHER: That's it. MOTHER: Mhm. CHILD: That's the way I can do it. MOTHER: You could probably tighten it with your hands first. MOTHER: It'll be a little easier. MOTHER: That's right. MOTHER: Good girl. CHILD: Now, do ___ store.

3	chi	18	the	MOTHER: Oh. CHILD: Cause I'm okay. MOTHER: Say you're pretty good at picking things up, aren't you? MOTHER: How come you're not that good at home at picking things up? CHILD: Well cause i have a lot of work to do. MOTHER: You have a lot of work to do. MOTHER: Oh. CHILD: Look what I have just for the baby. MOTHER: Mm. CHILD: For ___ baby.
3	chi	19	the	MOTHER: There we go. CHILD: You're gonna turn it. MOTHER: Your mommy's not very mechanical, I'm afraid. MOTHER: I'm not good at this stuff particularly. CHILD: How... how ... CHILD: Where's this... CHILD: How'd this go? MOTHER: It's the windshield, huh? CHILD: The windshield. CHILD: How do ya get ___ windshield back on?

3	chi	20	the	CHILD: Milk. MOTHER: Uhhuh bottles of milk. MOTHER: You think your family needs milk? MOTHER: Who are you? MOTHER: Are you the mommy or the daughter or the daddy? MOTHER: Who are you going shopping? CHILD: Mommy. MOTHER: You're the mommy? MOTHER: Who are you buying this food for? CHILD: For __ daddy and the sister.
3	chi	21	a	MOTHER: That's pretty silly, huh? CHILD: Yes. CHILD: Why? MOTHER: Okay, well Charlie, she stops to pick a flower. MOTHER: And what happens? CHILD: The baby gets out. MOTHER: He sure does. MOTHER: Look at how surprised she is! MOTHER: She's just amazed, isn't she? CHILD: Why is ... why he's hiding behind __ tree?

3	chi	22	a	CHILD: Can w ... can ... MOTHER: Let's ... let's ... MOTHER: No. MOTHER: Come sit down for a minute. MOTHER: Let ... let me finish this. CHILD: Okay. MOTHER: Who do you think this boy ... MOTHER: What is this? MOTHER: What is that? CHILD: __ bike ...
3	chi	23	a	MOTHER: Come help me find my kitty cat. CHILD: Okay. CHILD: Now ... now can I open up my store? MOTHER: Not yet. MOTHER: A few more minutes. CHILD: Okay. MOTHER: And what do you know? MOTHER: Do you know what that is? MOTHER: Hmm? CHILD: __ mailbox.

3	chi	24	a	CHILD: What's that? MOTHER: That's just a ... a ... a piano I think. MOTHER: I don't know. MOTHER: Can you read me the story though? MOTHER: Tell me what's in this book. CHILD: I wanna play the piano. MOTHER: Well. MOTHER: We will after we read the story. MOTHER: Let's do the story first. CHILD: Is this __ piano?
3	chi	25	a	MOTHER: We don't have hands. MOTHER: It makes it difficult. CHILD: I know. CHILD: Yes. CHILD: I know that. CHILD: It's hard, isn't it? MOTHER: Uhhuh. MOTHER: You just put that screwdriver right in that screw, didn't you? CHILD: Yes. CHILD: And that's __ black one.

3	chi	26	a	MOTHER: When you push these you ring up this, and what do you wanna have happen next? CHILD: Oh okay, one, two, three, six, seven, nine, ten, eleven, twelve, eight, that's twenty. MOTHER: Oh. CHILD: Thank you that's enough. MOTHER: Okay. CHILD: That's enough for me. MOTHER: That's enough for you? MOTHER: Thank you. CHILD: Okay now I'll get a bag. CHILD: I'll get ___ paper bag.
3	chi	27	a	MOTHER: Oh I thank you so much. CHILD: And would you like something to drink before you go? MOTHER: Oh yes. CHILD: And something to eat? MOTHER: Yes. MOTHER: I hear you're a very good cook. CHILD: See I'll cook you a ... MOTHER: What are you going to make for me? CHILD: I'm gonna make some ... I'm gonna make ... CHILD: Oh ___ spider.

3	chi	28	a	MOTHER: Oh payers don't put cars together? CHILD: No they don't. MOTHER: I didn't know that? MOTHER: How come? CHILD: I don't know. MOTHER: You don't know? CHILD: Hunhunh. MOTHER: But I don't know if I know how. MOTHER: Maybe you could help. CHILD: Okay can you ... can you give me __ tissue?
3	chi	29	a	MOTHER: Should I ... think I oughta get some coffee? CHILD: Oh I don't ... I don't know: if we have coffee maybe we do. MOTHER: Here, put these in cause I need those. CHILD: Here's a little sugar. MOTHER: Okay. MOTHER: Alright. MOTHER: Sugar, maybe we could use that. MOTHER: Napkins we're almost out of. CHILD: Well napkins we don't have. CHILD: I'll gonna give you some milk and __ can of soup.

3	chi	30	a	MOTHER: The man ... MOTHER: Well the branch wasn't strong enough to hold him, so he's gonna fall. CHILD: Oh. MOTHER: Thud, right on the ground. MOTHER: Well they're all kind of battered and bruised and they're tired of trying to find that cat. MOTHER: There he ... MOTHER: Look at him. MOTHER: He's walking along behind just as... like nothing happened. MOTHER: Proud as can be. CHILD: Is ... is he __ baby?
3	chi	31	a	MOTHER: Now. MOTHER: Do you want me to tell you a story? CHILD: Here. MOTHER: What's that? CHILD: A dime. MOTHER: No. MOTHER: A quarter. CHILD: Quarter? CHILD: I ... I'll put it back. CHILD: That camera's taking __ picture of me.

3	chi	32	a	CHILD: Okay. CHILD: Can I sit in this small chair? MOTHER: Yeah but let's play with it. CHILD: Okay. CHILD: Mom I want ... MOTHER: Are you gonna be Mr. Hooper? CHILD: I wanna play with this. MOTHER: Okay. MOTHER: What is that? CHILD: It's ___ typewriter.
3	chi	33	a	CHILD: Yeah. CHILD: Now... CHILD: Now ... MOTHER: Now what else do we hafta do? CHILD: We hafta put the tools in. MOTHER: What's that? CHILD: I don't know. MOTHER: Is that a wrench? MOTHER: And that's a screwdriver. CHILD: I know it's ___ screwdriver.

3	chi	34	a	MOTHER: You wanna drink of water? CHILD: Yeah. MOTHER: Oh, that's not gonna be too easy ...um, let me see. MOTHER: Why don't you wait just a minute until we look at the store ... at the ... at behind here, okay? MOTHER: And then we'll go down and get a drink, okay? MOTHER: Does this move? MOTHER: What'd you think you do with this? MOTHER: What'd you think you do with that? MOTHER: Hmm? CHILD: There's __ man?
3	chi	35	a	CHILD: Yep. MOTHER: Oh, look. CHILD: A piano. MOTHER: No, it isn't. MOTHER: It's a grocery store. CHILD: What's this? CHILD: See? MOTHER: We have our own grocery. MOTHER: What's that? CHILD: It's __ piano.

3	chi	36	a	CHILD: Yeah, I'm gonna drink the coffee. MOTHER: Well, here's good luck. CHILD: What? MOTHER: Good luck. CHILD: Good luck. MOTHER: Where's your coffee? MOTHER: You're not drinking. CHILD: No. CHILD: Not yet. CHILD: Here's ___ spoon.
3	chi	37	a	MOTHER: Oh do you think little Miss muffett's near by? CHILD: Yeah. MOTHER: He's a friendly spider. CHILD: Yeah. CHILD: Do you think he's trying to look for her? MOTHER: Maybe he's looking for little Miss muffet so he can sit next to her on the tuffet. CHILD: Mom? MOTHER: What? CHILD: Would ya get tissue ... CHILD: Do you have ___ tissue?

3	chi	38	a	MOTHER: The branch broke. CHILD: Yeah. MOTHER: The branch broke, and everybody fell down. MOTHER: And the kitty fell right into the mommy's hands. MOTHER: And they put the cowboy in the baby carriage. MOTHER: And the policeman needs a ... a stick to help him walk home. MOTHER: And the cat comes behind them. MOTHER: And the mommy goes in the house. MOTHER: And now what's she going to pretend to be? CHILD: ___ nurse.
3	chi	39	a	MOTHER: Hmm? CHILD: Look at that camera. MOTHER: Look at that camera. CHILD: What is it doing? MOTHER: It's taking a picture of you for posterity. CHILD: I want it to say.. CHILD: Click. MOTHER: Click? MOTHER: It's not that kind of camera, sweetheart. CHILD: Is it taking ___ picture of um.. of you?

3	chi	40	a	CHILD: Hah. MOTHER: Look at that. MOTHER: Can I look at the lights? MOTHER: Can you just pull on that? MOTHER: Does that just come apart? CHILD: Can't find this extra. CHILD: I can't tell'em. MOTHER: Oh! MOTHER: What's down in there? CHILD: ___ tire.
---	-----	----	---	--

D.3 Child determiner uses in 4-year-old corpora

age	speaker	item	ans	context
-----	---------	------	-----	---------

4	chi	1	the	MOTHER: But the cat landed in the arms of the little girl. MOTHER: There's a shoe in the air. MOTHER: And then they're all tired ... MOTHER: And the policeman's walking with a cane. MOTHER: He's so exasperated. MOTHER: He's been through so much, trying to catch the cat. MOTHER: And the cat's walking along behind like a little prince, huh? CHILD: Yeah. MOTHER: Yeah. CHILD: Turn __ page.
4	chi	2	the	MOTHER: In the water. MOTHER: And who's gonna go in the pipe? CHILD: The cat. MOTHER: And look it what is she saying? MOTHER: What is she saying to the policeman? CHILD: Get that cat. MOTHER: Get that cat because he's going back in the pipe. MOTHER: Oop the cat's in the pipe. MOTHER: And who's in the pipe now too? CHILD: __ girl.

4	chi	3	the	MOTHER: Hmm? MOTHER: You tell me? MOTHER: What happened? CHILD: The kitty cat. MOTHER: He got the kitty. MOTHER: That's right. MOTHER: And what's this boy doing here? MOTHER: Something not very nice. MOTHER: He's sticking his tongue out. CHILD: To __ policeman?
4	chi	4	the	MOTHER: Go ahead. CHILD: Why you need this? CHILD: You need this? MOTHER: Well that's just something to sit on. CHILD: When he gets done with the VA then what he's gonna do? MOTHER: Who? CHILD: He's gonna come in the school? MOTHER: Who? MOTHER: Who was? CHILD: __ daddy.

4	chi	5	the	MOTHER: There. MOTHER: Okay, now I think you can screw that back in. MOTHER: Turn it the other way. MOTHER: Turn it ... MOTHER: That's right. CHILD: You go this way? MOTHER: Yeah. MOTHER: You turn it in one direction when you wanna take the screw out. MOTHER: And you turn it in the other direction when you wanna put it in. CHILD: Is this __ way that you hafta put it in?
4	chi	6	the	MOTHER: We're not reading it. MOTHER: We're just making up a story. MOTHER: Do you ever dress up like that little girl? CHILD: No. MOTHER: You don't? CHILD: I wanna play with this. CHILD: Cmon. MOTHER: What happened here, Helen? MOTHER: What happened here? CHILD: __ cat ran out of the carriage and ... and ... and ... and the cat went away.

4	chi	7	the	MOTHER: Helen? CHILD: No. MOTHER: You're no ... you're sure? MOTHER: Look at this. CHILD: No. MOTHER: Now what happened? CHILD: Now the cat ran away. MOTHER: Again? MOTHER: And what happened over here? CHILD: ___ policeman fell down.
4	chi	8	the	MOTHER: Now. MOTHER: What's gonna be number six? CHILD: Number six. CHILD: Gonna be hard to do that. MOTHER: Not really. MOTHER: I bet if you try you can do it. MOTHER: You hafta do it with the screw. MOTHER: See these are screwed in again. MOTHER: It's the same type. CHILD: So you hafta screw it with ___ screwdriver.

4	chi	9	the	CHILD: Say. CHILD: Number thirteen. CHILD: Thirteen? MOTHER: Okay. MOTHER: Number thirteen. CHILD: Then we'll... MOTHER: Number ... MOTHER: Fourteen! CHILD: Say number six ... CHILD: Say __ number ...
4	chi	10	the	CHILD: Yeah. MOTHER: I think they're trying to get the cat to come. CHILD: Right. MOTHER: Over to them. CHILD: Yeah. MOTHER: And the cat doesn't seem to wanna.. MOTHER: Do you think that's right? CHILD: Right. CHILD: And ... and now ... now he ... now he's sitting on there. CHILD: And ... and he trying to get __ cat.

4	chi	11	the	CHILD: And then .. MOTHER: No. MOTHER: Wait. MOTHER: We didn't do this page yet. CHILD: This page? MOTHER: Yeah. CHILD: This page? MOTHER: We didn't do either of these pages yet. CHILD: Um ... um and ... and they're splashing into the water. CHILD: And um ... and they're drowning because ____ cat's going in the hole.
4	chi	12	the	CHILD: Yeah. MOTHER: Oh. CHILD: And ... and this ... CHILD: Now he's climbing up um .. there ... and he has messy hair. CHILD: And he doesn't like it. CHILD: So he's trying to catch that. MOTHER: I see. MOTHER: What a complicated story. MOTHER: Isn't it? CHILD: And now ... now he's climbing up on ____ tree.

4	chi	13	the	CHILD: Put the screwdriver away. CHILD: Put the jack in there. MOTHER: Didn't you forget something? CHILD: What? MOTHER: How can we drive this car? CHILD: Would ya ... MOTHER: This is a great car, isn't it? CHILD: There we go. CHILD: Put the wrench away. CHILD: Put the other screwdriver in with __ jack.
4	chi	14	the	CHILD: I'll tell you when! MOTHER: I know but you're so slow. MOTHER: I'm really in a rush. CHILD: Okay. CHILD: Mom. MOTHER: Yes? CHILD: Sometimes you tell me where things are, okay? MOTHER: What do you mean? CHILD: Like when you can't ... CHILD: When you don't see things on __ shelf.

4	chi	15	the	MOTHER: It looks like he's really in the quicksand. MOTHER: Oh and they both go in head first, or is that water? MOTHER: It must be water. MOTHER: And so that ... when they when they're stuck down there, the cat goes back the other way. CHILD: They're gonna drowned. MOTHER: Mhm. MOTHER: Oh and look what happened here. MOTHER: The little ... the little boy waited on the other side, and he's the one who caught the cat. MOTHER: Oh, it's ... CHILD: Is __ man angry?
4	chi	16	the	CHILD: No, Mom. CHILD: I gotta keep the front open. MOTHER: Okay, is there something wrong with the engine? CHILD: I don't know. MOTHER: Oh. CHILD: I gotta get it ... get this out. CHILD: There is something wrong with ... CHILD: Now there is something wrong with this engine. MOTHER: What's the matter with it? CHILD: Maybe__ car is stuck in the mud.

4	chi	17	the	MOTHER: Do you want me to help you you mean? CHILD: Yeah. MOTHER: So you can get done faster? CHILD: I ... I ... I ... I'll ... I do this many things ... CHILD: I ... I do this many. CHILD: And you do this many. MOTHER: Okay. MOTHER: Which ones do you want me to take apart? MOTHER: You tell me which ones. CHILD: Do ... do ___ wall.
4	chi	18	the	CHILD: Hunhunh. CHILD: Hey. CHILD: What do ... what do they want us to do? MOTHER: I think they want us to sit here and try and put the car back together. MOTHER: See if we can. CHILD: One time the lady did because ... because I couldn't. MOTHER: Okay. MOTHER: Come back on here Davey. MOTHER: And sit down. CHILD: One time ___ lady did it.

4	chi	19	the	CHILD: What? MOTHER: The great cat chase. CHILD: Mommy, what's in there? CHILD: Nothing? MOTHER: Nothing. CHILD: I wanna play with the car. MOTHER: Well, let's read the story. CHILD: No. CHILD: I wanna ... CHILD: No, I play with ___ cash register and you read the story to me while I do it.
4	chi	20	the	MOTHER: Where's the cat Edward? MOTHER: On the outside. MOTHER: Who's going inside? CHILD: Policeman. MOTHER: Policeman. MOTHER: Where's the boy? MOTHER: On top right. MOTHER: I wonder what's going to happen next. MOTHER: What do you think is going to happen? CHILD: ___ cat's going to run away.

4	chi	21	a	MOTHER: Oh. MOTHER: We've got one more wheel back here before you can take the sides out and ... and other screws here. CHILD: I know. MOTHER: Want me to do this one? MOTHER: Or do you wanna do it? CHILD: Do this one. CHILD: And I'll take ... CHILD: Let me take the screws off. MOTHER: You know what this part of the car's called? CHILD: It's only ___ screw.
4	chi	22	a	MOTHER: It's a screw on this car. MOTHER: But on real cars it's called a hubcap. CHILD: This is a hubcap. MOTHER: Well. MOTHER: It's not really a hubcap. MOTHER: But it looks like the hubcap part of a real car. CHILD: It's really just a screw. MOTHER: Yeah. CHILD: It's not a screw. CHILD: It's ... it's ___ metal.

4	chi	23	a	MOTHER: The one who's in charge. CHILD: Hi. CHILD: I'm Mr man eating fish. CHILD: And I'm wanna catch you. MOTHER: Alright. MOTHER: I think you'd better get down. CHILD: Oh. MOTHER: Uhh oh. MOTHER: Oh. CHILD: You are __ baby.
4	chi	24	a	MOTHER: Alright. MOTHER: Okay we'll play. MOTHER: Would you like to play first? MOTHER: Alright ... what's this? MOTHER: Oh my you don't have this one, do you? CHILD: No. MOTHER: How about this? CHILD: Alright. MOTHER: What is it? CHILD: __ toy car.

4	chi	25	a	MOTHER: Where's the cat? CHILD: Under there. MOTHER: Where is the cat? MOTHER: No. CHILD: Right there. MOTHER: Right it's hiding. MOTHER: What's, going to happen next I wonder. MOTHER: She cries and she calls who? MOTHER: Who's this? CHILD: __ policeman.
4	chi	26	a	MOTHER: Would you like a cookie? CHILD: Yeah. MOTHER: Here. CHILD: Thank you, mommy. MOTHER: Oh, you're welcome, Helen. MOTHER: How about ... MOTHER: Let's see, oh look what's here, Helen. CHILD: What? MOTHER: Some baby cereal. CHILD: Here's __ pound of money.

4	chi	27	a	MOTHER: Who popped out of the carriage? MOTHER: Ooh look what she's doing. MOTHER: Ooh where's my baby? MOTHER: Where is he? CHILD: Here. MOTHER: And she's rushing along. MOTHER: Where is that? MOTHER: Where's my baby? MOTHER: I've lost my baby who's this? CHILD: __ policewoman.
4	chi	28	a	MOTHER: A rain barrel it looks like. MOTHER: Yeah. MOTHER: Like a big pipe. MOTHER: The cat looks pretty angry. MOTHER: Uhh oh. MOTHER: Somebody's crawling in. MOTHER: The policewoman is crawling in. MOTHER: And look who's out. MOTHER: But what is this over here? CHILD: That's __ lake.

4	chi	29	a	MOTHER: Oh. MOTHER: Okay. MOTHER: Let's see. CHILD: Okay. CHILD: Now. CHILD: And you know what? CHILD: And now ... now she ... um ... now she ... MOTHER: What? CHILD: Umm looking for help. CHILD: And ___ policeman's blowing his whistle because ... because um ... the cat was ... because the cat ... cause the cat was looking ... cause the cat was in the stroller.
4	chi	30	a	CHILD: Sherry's was different. MOTHER: Sherry was ... Sherry's was different? CHILD: Yeah. MOTHER: You mean the study she was in? CHILD: Yeah. MOTHER: Yeah. CHILD: Because I ... cause I have a truck and stuff. MOTHER: Right. CHILD: And I got a store. CHILD: And I ... and I got ___ book.

4	chi	31	a	CHILD: Not funny. CHILD: Dad didn't even see that picture. MOTHER: Dad didn't see that one? CHILD: Mhm. MOTHER: Oh. CHILD: He missed it. MOTHER: He missed that one? MOTHER: Gee, now the cat's coming to this tunnel. MOTHER: Yep. CHILD: It's not __ tunnel.
4	chi	32	a	MOTHER: Huh? MOTHER: I'd like to see what's in it. MOTHER: Thank you. MOTHER: How about if we close up shop and read the story? CHILD: Yeah. MOTHER: Alright? MOTHER: Here. MOTHER: We just played make believe so you can put the money back in there. CHILD: Wow! CHILD: __ dime came out.

4	chi	33	a	MOTHER: Or do you wanna do it? CHILD: Do this one. CHILD: And I'll take ... CHILD: Let me take the screws off. MOTHER: You know what this part of the car's called? CHILD: It's only a screw. MOTHER: Well. MOTHER: It's a screw on this car. MOTHER: But on real cars it's called a hubcap. CHILD: This is ___ hubcap.
4	chi	34	a	CHILD: Then you will teach me ... MOTHER: Keep turning ... keep turning, keep turning, just move your wrist as you turn. MOTHER: Can you move your wrist? MOTHER: Like this alright. MOTHER: Oh it's almost off. MOTHER: It's almost off. MOTHER: That's like it's at his garage. MOTHER: You know when I get a flat tire? MOTHER: Remember we went there and had our tire fixed? CHILD: Yeah this one has ___ flat tire.

4	chi	35	a	CHILD: Then I could play with it? MOTHER: Later, okay? CHILD: With it? CHILD: Then I'm going to play with the car after the story. MOTHER: Okay, let's ... let's see if we can make up a story here. CHILD: Okay. MOTHER: Alright? MOTHER: There. MOTHER: What's going on here? CHILD: The cat is in ___ carriage.
---	-----	----	---	--

4	chi	36	a	CHILD: And ... and then they were looking and they ... and then she ... she thought that they were ... different umm ... from their cat. CHILD: So she ... now she ... now all of them didn't know that there was that kind of cat in there. MOTHER: You mean they thought it was a different cat? CHILD: Yeah. MOTHER: What kind of cat did they think it was? MOTHER: What's the difference between that kind of cat and this kind of cat? CHILD: This kind ... CHILD: See. CHILD: They were ... CHILD: She ... she ... she thought that ... that there was ___ baby.
---	-----	----	---	---

4	chi	37	an	CHILD: Look at them. MOTHER: I'm not sure if this screw comes apart. MOTHER: Hmpf. MOTHER: Let's see if we can use something. MOTHER: But I don't know, Terry. CHILD: This. MOTHER: Do you think ... MOTHER: Oh. MOTHER: Hmm. CHILD: I got ___ idea.
4	chi	38	a	MOTHER: Can you pack that up for me? MOTHER: Cause I'm really in a hurry. CHILD: Okay. MOTHER: I'm waiting for the poor cash register there. CHILD: There. CHILD: And ... MOTHER: Okay. CHILD: Put ... put them ... MOTHER: Thank you. CHILD: ___ bag.

4	chi	39	a	MOTHER: Alright, can the ... does the kitty cat see her? CHILD: Uhhuh. MOTHER: Okay, now, what's happening over here? MOTHER: Who's that? CHILD: Policeman. MOTHER: Is that really or do you think it's two little kids playing make believe. CHILD: Two little kids making believe. MOTHER: Yeah. MOTHER: Well what's this boy? CHILD: Umm riding __ bike.
4	chi	40	a	MOTHER: Wow. CHILD: Mm mm mm. MOTHER: Is that very high? CHILD: Yep. MOTHER: Oh wow. MOTHER: What happens when the policeman falls down? CHILD: He hurts himself. MOTHER: Yeah but something fell off of him. MOTHER: What fell off of him? CHILD: __ shoe.

D.4 Mother determiner uses in 2-year-old corpora

age	speaker	item	ans	context
2	mot	1	the	MOTHER: Who's that? MOTHER: Is he driving the car? CHILD: Help me do this, mom. MOTHER: Okay. MOTHER: You have turn the wheel hard. MOTHER: It's just like Wendy's little man. MOTHER: It comes off. MOTHER: Oop. MOTHER: See? MOTHER: But how can you drive __ car now?
2	mot	2	the	CHILD: Yeah. MOTHER: Okay. CHILD: We'll cook dinner out here. CHILD: What's in here? MOTHER: What's in there? CHILD: Yeah. MOTHER: It's empty. MOTHER: Want me to read the ... MOTHER: Let's see ... MOTHER: What's on __ box?

2	mot	3	the	MOTHER: Look at that. MOTHER: Way up the tree. CHILD: Let ... let ... let me see. MOTHER: Oh! MOTHER: My goodness! CHILD: What? MOTHER: The policeman went after him. MOTHER: And cats are very good at climbing trees. MOTHER: But sometimes policeman aren't. MOTHER: And ___ cat gets on this little bitty branch.
2	mot	4	the	CHILD: Yeah. MOTHER: Let me see. MOTHER: That's the wrong way. MOTHER: You wanna turn the other way that's right. CHILD: Okay. MOTHER: That's very good. CHILD: Where's my sweater? MOTHER: We ... your sweater ... we ... we took your sweater ... MOTHER: Some ... MOTHER: ___ lady took your sweater in the other room with my raincoat so that it'll be safe.

2	mot	5	the	MOTHER: Yeah. CHILD: Where does this go? MOTHER: Turn it over, and you see if you can find the place. CHILD: Another one. CHILD: Another one. MOTHER: You know what this is called? CHILD: Why? MOTHER: That's called a screw. CHILD: These are screws? MOTHER: Well, where we've put in ___ screwdriver and take it out, it's called a screw.
---	-----	---	-----	--

2	mot	6	the	MOTHER: And then after we play ... we do the things that we're sposta do, we can then take a look at the whole room, okay? MOTHER: Do you wanna play with this little car? CHILD: Yeah. MOTHER: Okay, have you seen a car like this before? CHILD: Yeah. MOTHER: Oh, you know what they call this? MOTHER: They call this a convertible because it doesn't have a top. MOTHER: See there's no top on it? MOTHER: Those are tools. MOTHER: To fix ___ car if it gets broken.
2	mot	7	the	CHILD: Okay. MOTHER: Here. CHILD: I'll take it out. MOTHER: Alright well, let's move these off. MOTHER: Take these off. MOTHER: Move the tools off the truck, okay? MOTHER: Now, let's pick it up. MOTHER: But give me the nut and bolt. MOTHER: First hand me the ... hand me the bolt. MOTHER: ___ red bolt, that's right.

2	mot	8	the	CHILD: Oh. MOTHER: Now you wanna try the other one? MOTHER: You try this one. CHILD: Where? MOTHER: Okay? MOTHER: Here's the tire. CHILD: Okay. MOTHER: That's right. CHILD: Take this. MOTHER: Here, I'll hold ___ tire in place for you, okay?
2	mot	9	the	MOTHER: Here, I'll hold the tire in place for you, okay? MOTHER: Now, fir ... MOTHER: No. MOTHER: First ... MOTHER: What do you hafta do first? MOTHER: Is that right? MOTHER: What do you hafta do first? MOTHER: No. MOTHER: Then it ... the bolt goes in this way. MOTHER: But ___ nut goes in here.

2	mot	10	the	MOTHER: Or you wanna read the story? CHILD: You read the story. MOTHER: Okay. MOTHER: Once upon a time ... there was a lady ... and a kitten. MOTHER: Here's a cat. MOTHER: Where do you see the cat? MOTHER: See? MOTHER: The cat's in the carriage. MOTHER: Isn't that silly? MOTHER: One day ___ lady took her cat for a walk.
2	mot	11	the	MOTHER: There they go, plunk. MOTHER: They're falling. MOTHER: They're falling. MOTHER: Oh. MOTHER: No. MOTHER: The policeman fell right on top of the little boy. MOTHER: But the goosey fell right into the lady's arms. MOTHER: And here they go. MOTHER: The goosey, the policeman, the lady, and the little boy ... MOTHER: Look at ___ little boy in the carriage.

2	mot	12	the	MOTHER: There's ... there's some grass growing in the water. MOTHER: All this is water. MOTHER: The policeman fell in the water. MOTHER: And here comes the lady. MOTHER: And I think the lady's gonna fall in the water. MOTHER: And look at Googey. MOTHER: Googey's crawling up the mountain. CHILD: Look at Googey. MOTHER: Yeah! MOTHER: He's crawling up ___ mountain and down the side.
2	mot	13	the	MOTHER: They sure did. MOTHER: Real wet. CHILD: Did they get hurt? CHILD: Did she get her hair wet? MOTHER: Yeah. MOTHER: Doesn't it look like it? MOTHER: And even her hat fell off. MOTHER: It's on the policeman's head. MOTHER: See? MOTHER: Don't they look funny in ___ puddle?

2	mot	14	the	CHILD: He fell off. CHILD: Off of the tree. CHILD: Why did he? MOTHER: Well. MOTHER: The branch broke. MOTHER: See? MOTHER: The branch broke. MOTHER: And the policeman fell on top of him. MOTHER: See? MOTHER: ___ policeman was too heavy for the branch.
2	mot	15	the	MOTHER: And .. I'm gonna buy some oreos too, okay? MOTHER: It's empty. MOTHER: Okay? MOTHER: Okay? MOTHER: And I'll give you the money. MOTHER: And will you put it all in a bag for me so I can take it home? CHILD: Yeah. MOTHER: Okay. MOTHER: Here is the money, okay? MOTHER: You wanna put it in ___ cash register?

2	mot	16	the	MOTHER: Can I help? MOTHER: You hafta line it up. MOTHER: And it's very hard to do. MOTHER: Why don't I line it up, and you push it. MOTHER: Go ahead. MOTHER: Push. CHILD: Uh. MOTHER: There you go. MOTHER: Let's put the screwdriver. MOTHER: Can you put ___ screwdriver back?
2	mot	17	the	CHILD: Why? MOTHER: Are they going swimming? CHILD: They going swimming. MOTHER: In a pool? CHILD: Yeah. MOTHER: Or in the water? CHILD: In the water. MOTHER: In a lake. CHILD: In a lake. MOTHER: And ... and who's going back in ___ pipe?

2	mot	18	the	MOTHER: Very good. CHILD: Oh! CHILD: How about that? MOTHER: Yeah. MOTHER: Let me put that away too. MOTHER: Here we go. MOTHER: Look at this. CHILD: Oh. MOTHER: Look at that. MOTHER: Look at ___ lady.
2	mot	19	the	CHILD: This the couch? MOTHER: That's what it looks like. CHILD: Ohh. CHILD: I put ... I put it over here? MOTHER: Mhm. CHILD: What's in here? MOTHER: Well, I think they're all empty. MOTHER: It's just like a ... MOTHER: It's a pretend store. MOTHER: You know when you go to ___ store?

2	mot	20	the	MOTHER: And then ... and then we can go shopping and when we come home we can eat, since you set the table already, okay? CHILD: This be yours and this be somebody else's. MOTHER: Okay, thank you. CHILD: What's ... what's the other one in there? MOTHER: Here. MOTHER: Do you know what this is? MOTHER: This is play money. CHILD: Oh. MOTHER: Okay? MOTHER: You wanna stay by ___ cash register.
2	mot	21	a	CHILD: We need screw ... MOTHER: Here is a screwdriver. MOTHER: But you need a screw. MOTHER: Where's your screw, Laurie? MOTHER: Do you see a screw here on the floor? MOTHER: That's a screwdriver. MOTHER: Do you see a screw? CHILD: Hunhunh. MOTHER: No. MOTHER: This's ___ screw.

2	mot	22	a	MOTHER: Yeah. CHILD: Where does this go? MOTHER: Turn it over, and you see if you can find the place. CHILD: Another one. CHILD: Another one. MOTHER: You know what this is called? CHILD: Why? MOTHER: That's called a screw. CHILD: These are screws? MOTHER: Well, where we've put in the screwdriver and take it out it's called ___ screw.
2	mot	23	a	CHILD: Why? MOTHER: That's called a screw. CHILD: These are screws? MOTHER: Well, where we've put in the screwdriver and take it out it's called a screw. CHILD: Uh, I can't get it in. MOTHER: The ... you know what this is called? CHILD: Nah. MOTHER: What's that part of the car called? MOTHER: That part that turns? MOTHER: ___ wheel.

2	mot	24	a	MOTHER: And what's this? CHILD: A quarter. MOTHER: No, this is a quarter. MOTHER: This is a nickel, a quarter, a nickel, a dime and a ... MOTHER: What's this? CHILD: A penny. MOTHER: Penny. MOTHER: Which is the biggest one? MOTHER: Look here. MOTHER: Oh, here's ___ dollar okay, now I have a dollar.
2	mot	25	a	CHILD: No. MOTHER: Can you count these bottles? CHILD: One, two. MOTHER: One, two ... CHILD: One, two, three, four, five. MOTHER: Oh, I think you have an extra one. MOTHER: Count again. CHILD: And I have a extra one. CHILD: I have an extra one. MOTHER: You have ___ extra one?

2	mot	26	a	MOTHER: Okay? MOTHER: That's a picture. MOTHER: We can see them from over here. MOTHER: And then after we play ... we do the things that we're sposta do, we can then take a look at the whole room, okay? MOTHER: Do you wanna play with this little car? CHILD: Yeah. MOTHER: Okay, have you seen a car like this before? CHILD: Yeah. MOTHER: Oh, you know what they call this? MOTHER: They call this a convertible because it doesn't have __ top.
---	-----	----	---	---

2	mot	27	a	CHILD: Yeah. MOTHER: It's just like it only it's strawberry instead. MOTHER: Oh, here. MOTHER: Here's one that's ... MOTHER: What's that kind? CHILD: Chocolate. MOTHER: Mhm. MOTHER: Do you like chocolate or strawberry better? CHILD: Strawberry. MOTHER: We'll hafta tell grandma that they make __ strawberry kind, and maybe she'll get it when you got to visit.
2	mot	28	a	MOTHER: Giddyay ... CHILD: Giddyay ... MOTHER: Giddyay up? CHILD: Yeah. MOTHER: Yeah. CHILD: Look. MOTHER: What's that? CHILD: That's the girl. MOTHER: Come over here. MOTHER: That's not __ girl, Richard.

2	mot	29	a	CHILD: Sit on the chair please. CHILD: Let's look something. MOTHER: Okay. MOTHER: I'm sitting up. MOTHER: You wanna read me the story? MOTHER: Or you wanna read the story? CHILD: You read the story. MOTHER: Okay. MOTHER: Once upon a time ... there was a lady .. and a kitten. MOTHER: Here's ___ cat.
2	mot	30	a	CHILD: Three hundred dollars. MOTHER: Three hundred dollars? CHILD: Yeah. MOTHER: Oh, I don't have three hundred dollars. MOTHER: Will you give it to me for free? MOTHER: No? MOTHER: What kind of money is that? MOTHER: Do you know what that's called? CHILD: Cash register. MOTHER: Yeah well, this is ___ cash register.

2	mot	31	a	CHILD: Right here? CHILD: Oh. CHILD: Right here. MOTHER: Well ... Close. MOTHER: Bring the car back over here. MOTHER: And I'll show you. MOTHER: I'll help you with it. CHILD: Okay. MOTHER: But bring back the other pieces also. MOTHER: You know what ___ screwdriver does Richard?
2	mot	32	a	MOTHER: You know what a screwdriver does Richard? MOTHER: What is it? CHILD: No, nothing. MOTHER: Pretend? CHILD: What's in here? CHILD: Open it. MOTHER: You know, these don't open, they're for pretending with. MOTHER: What does that look like? CHILD: I don't know. MOTHER: ___ can of what?

2	mot	33	a	CHILD: Yeah. CHILD: You take the ... CHILD: Are the blue ones out? MOTHER: Oh there's tools in here. MOTHER: Actually you can do this with your fingers. MOTHER: You do it. MOTHER: Oops! MOTHER: A fender came off? CHILD: What, mama? MOTHER: ___ fender fell off?
2	mot	34	a	CHILD: And ... and some back here. MOTHER: Some back here. MOTHER: Mhm. CHILD: And some there? MOTHER: Mhm. MOTHER: And the price. MOTHER: How much it cost. MOTHER: Okay, let's see what happens. MOTHER: What's this? MOTHER: Looks like a little girl dressed up .. taking somebody for a ride In ___ carriage.

2	mot	35	a	MOTHER: What if you do this? MOTHER: Slide it right in there. MOTHER: See if it fits. CHILD: Here? MOTHER: Right in here. MOTHER: Good. MOTHER: Get it all in? MOTHER: There. MOTHER: And now we hafta put something in there to hold it. MOTHER: How bout ___ screw?
2	mot	36	a	CHILD: No. CHILD: We hafta get ... CHILD: Oh. CHILD: They fell off. MOTHER: Oh, Laurie. MOTHER: We must not have put them on tight enough. MOTHER: You know? CHILD: Would you do this again? CHILD: Would you ... MOTHER: ... Like me to read you ___ book?

2	mot	37	a	MOTHER: Oh, Laurie. CHILD: Hmm? MOTHER: Look at this book. MOTHER: Let's sit down right here and read, okay? MOTHER: There we go, okay? MOTHER: This is about a cat. MOTHER: Oh. MOTHER: Look. MOTHER: The little girl is taking her cat for a walk. MOTHER: She has him in ___ baby carriage.
2	mot	38	a	MOTHER: What does she have for them? MOTHER: Yeah. MOTHER: What ... what is she eating? MOTHER: Cookies and juice? CHILD: Yeah. MOTHER: And what is she dressed like? MOTHER: She's not dressed like a mommy anymore. MOTHER: What's she dressed like? MOTHER: Is she dressed like a nurse? MOTHER: She is ___ nurse.

2	mot	39	a	CHILD: Why? MOTHER: Why? MOTHER: Do you know what this is called? CHILD: What? MOTHER: This is called a baby carriage. MOTHER: Where they put ... is that a baby in the carriage? MOTHER: Look closely huh. CHILD: I think you're in it! MOTHER: He's gone for a ride. MOTHER: But is that ___ baby?
2	mot	40	a	MOTHER: Who does that little boy look like? MOTHER: He's wearing cowboy boots. CHILD: He's wearing cowboy boots. MOTHER: Yeah, he looks like Andrew. MOTHER: What's he doing? CHILD: Where's that other one? MOTHER: Uh, I don't know. MOTHER: Here. MOTHER: I think that's a jack. MOTHER: He's riding ___ bike, anyway.

D.5 Mother determiner uses in 3-year-old corpora

age	speaker	item	ans	context
3	mot	1	the	MOTHER: Oh! MOTHER: Who's in the carriage? CHILD: The little boy. MOTHER: And you know what? MOTHER: If you got in Rachel's carriage you'd look like that because you'd be too what? CHILD: Big. MOTHER: Too big for it. MOTHER: That's right. MOTHER: And look at the little boy. MOTHER: And ___ kitty.
3	mot	2	the	MOTHER: Oh. CHILD: Yeah. MOTHER: You're right. MOTHER: His shoe fell off, didn't it? CHILD: Yes. MOTHER: Yeah. MOTHER: And what happened here? CHILD: What? CHILD: What? MOTHER: ___ branch broke.

3	mot	3	the	MOTHER: Can I ... can I buy something at your store? MOTHER: And I'll pay you for it? CHILD: Yeah. MOTHER: Okay, I wanna buy some coffee. MOTHER: Look. CHILD: What? MOTHER: Can I buy this coffee? CHILD: Look it. MOTHER: A penny. MOTHER: how much is ___ coffee ?
3	mot	4	the	MOTHER: Should we put that one back in too? CHILD: Yup. MOTHER: Okay just right. MOTHER: And what holds that on? CHILD: That. MOTHER: Can you find it behind you? CHILD: Hey what about these stuff? MOTHER: What's behind you? MOTHER: Where's the one that fi... holds that on? MOTHER: We gotta have something to hold on ___ engine, though.

3	mot	5	the	MOTHER: But didn't he trip on his whistle? MOTHER: Did he trip on his whistle? MOTHER: He did. MOTHER: And he went to catch. MOTHER: And he says... MOTHER: "Don't worry." MOTHER: He says ... MOTHER: "I'll get your kitty for you." MOTHER: See? MOTHER: So he went to get ___ kitty.
3	mot	6	the	MOTHER: And you know what he's going do? CHILD: What? MOTHER: He's gonna go back inside the pipe and run to the other end. CHILD: How about this page? MOTHER: And now they're gonna go back inside too and look for him. MOTHER: And what do ya think happened? MOTHER: Who found the kitty? CHILD: The little boy. MOTHER: Yeah. MOTHER: And what does ___ lady ...

3	mot	7	the	CHILD: Hippo! CHILD: Hippo! MOTHER: Hippo? MOTHER: That's not a hippo, is it? MOTHER: Yes it is. MOTHER: You're right. MOTHER: You're right, Katie. MOTHER: How did you know? CHILD: See the man. MOTHER: ___ man is taking your picture.
3	mot	8	the	MOTHER: Oh. MOTHER: But the mommy ... MOTHER: Look what she said. MOTHER: Oh. MOTHER: I'm so happy that you got my kitty back. MOTHER: Ya see? MOTHER: See her saying that? MOTHER: Anyway. CHILD: What happened? MOTHER: Just as he was bringing ___ kitty down to the mommy didn't the branch break?

3	mot	9	the	CHILD: This is my room. MOTHER: Alright. CHILD: And where's this? MOTHER: That's umm ... CHILD: What? MOTHER: Umm what is it? CHILD: What? MOTHER: That's the kitchen. MOTHER: Now. MOTHER: This is __ kitchen.
3	mot	10	the	MOTHER: Turn it around. MOTHER: Turn it around. MOTHER: Okay. MOTHER: You can put your men in. MOTHER: Put the men in, Olivia. MOTHER: Okay? CHILD: No. CHILD: I can't. MOTHER: And we'll put this in. MOTHER: And we'll drive __ car around.

3	mot	11	the	CHILD: Not there anymore. MOTHER: Nope. MOTHER: He's all gone. CHILD: I got it in the tissue. MOTHER: Well no. MOTHER: He's not in the tissue. MOTHER: He's right there. MOTHER: Do you want me to get rid of him? CHILD: Yes. MOTHER: You don't like ___ spider?
---	-----	----	-----	---

3	mot	12	the	MOTHER: Does it hurt? CHILD: No. MOTHER: Okay now the pussy cat begins to go into a big tunnel here. MOTHER: And here comes Mathilda and Charlie following. MOTHER: And Craig's just following along to see what happens. MOTHER: There goes Charlie. MOTHER: He jogs into the tunnel too to following the pussy cat. MOTHER: And see? MOTHER: This tunnel goes through a big mound of the dirt. MOTHER: And there's ___ pussycat standing on top of the tunnel.
---	-----	----	-----	--

3	mot	13	the	MOTHER: Oh I'm sorry. MOTHER: Umm and here's Mathilda ever so happy that Charlie got the cat. MOTHER: But Craig isn't so ha... happy any more cause he thought he was going to be able to get it. MOTHER: Just then what happens? MOTHER: Look it there. MOTHER: The branch begins to break and Charlie drops the pussycat right into the arms of Mathilda. MOTHER: Now Mathilda has the pussy cat. MOTHER: But poor Charlie has landed right on top of Craig. MOTHER: Ah and look who ends up in the baby carriage. MOTHER: Not ___ pussy cat but Craig because poor Charlie landed on top of Craig, and look at poor Charlie.
---	-----	----	-----	---

3	mot	14	the	MOTHER: You give me some change. MOTHER: Okay? CHILD: Okay, maybe I'll give you the same ones. MOTHER: Okay. MOTHER: Thank you. CHILD: Now give'em all to me. MOTHER: Now give them all back to you? MOTHER: Okay let's put these back on the shelf, and let's read our story. CHILD: No I'm not gonna read a story today. MOTHER: Oh we hafta read __ story.
3	mot	15	the	MOTHER: A pitcher of lemonade. MOTHER: And everybody's happy, even the cat. MOTHER: You wanna read it again? CHILD: Yeah. MOTHER: Okay? MOTHER: Alright, now that we've seen the story one time? MOTHER: Okay okay okay, that's a good book. MOTHER: I think that this person makes really nice books. CHILD: Oh. MOTHER: Okay, you wanna tell me __ story this time?

3	mot	16	the	MOTHER: Okay. CHILD: Some more in. MOTHER: Can I have some milk? MOTHER: Please? MOTHER: And some sugar please? CHILD: And where's oh where's the sugar? CHILD: Where? MOTHER: Let me see if we can have sugar. MOTHER: Do you remember.. MOTHER: We bought sugar at ___ store?
3	mot	17	the	CHILD: What? MOTHER: What's that? MOTHER: Do you know? CHILD: What? MOTHER: They're in the water. MOTHER: At least it looks like water to me. MOTHER: I think that's what they're in. MOTHER: And he's helping her get out. MOTHER: And they fall straight in the water altogether, don't they? MOTHER: Look how happy ___ cat looks.

3	mot	18	the	CHILD: You too. MOTHER: Okay. MOTHER: Okay. CHILD: And these would be the guys who owned it. MOTHER: And those will be the guys who owned it? MOTHER: Okay, what're they gonna be doing while we're being mechanics? CHILD: Waiting. MOTHER: They're gonna wait. CHILD: Right down here. MOTHER: Oh okay, they're not going to take __ bus back from Watertown Square, are they?
3	mot	19	the	MOTHER: If you want. CHILD: Hah. MOTHER: You know what? CHILD: What? MOTHER: You know what we could do? MOTHER: We could pretend that we have a camera and take a picture of the camera. CHILD: What camera? MOTHER: Let's see. MOTHER: Let's pretend ... let's pretend that this is a camera. MOTHER: And you take a picture of __ camera.

3	mot	20	the	MOTHER: Okay, you got everything in your grocery cart? MOTHER: Okay. CHILD: Uhhuh. MOTHER: Are you watching where you are going? MOTHER: Uh oh! MOTHER: You almost knocked right into the cashier. MOTHER: Okay, here we are. MOTHER: You're at the cashier's place. MOTHER: I've got to get a bag for you to put your grocery in, right? MOTHER: Cause that's how they do it at ___ grocery store.
3	mot	21	a	MOTHER: And you hafta get them for me now. MOTHER: I'll ask ya... hafta ask you for them first. MOTHER: And hello. MOTHER: Excuse me. MOTHER: Can I have some groceries please? MOTHER: I'd like some corn flakes. MOTHER: I'd like a box of corn flakes. MOTHER: Right. CHILD: Here. MOTHER: I'd like ... ___ box of ah tea ... tea bags.

3	mot	22	a	CHILD: You sure? MOTHER: Yeah, you know what, Charlie? MOTHER: Go at it from the front. MOTHER: Go ... go this way and then do it. MOTHER: I think that'll make it easier. CHILD: How are you ... CHILD: Hey it's stuck! MOTHER: Look, you hafta turn one way or the other way until it gives a little bit to see which way it goes. MOTHER: There you go. MOTHER: That's ___ boy.
3	mot	23	a	MOTHER: Alright, so let me put down my stuff. MOTHER: Did you hear what Janet was telling you? CHILD: Which? MOTHER: Huh? MOTHER: She was telling me. MOTHER: Well... MOTHER: There are three different things to do, okay? MOTHER: There's a book to look at. MOTHER: There's a take-apart car. MOTHER: And there's ___ store, sort of like Josh has at his house.

3	mot	24	a	CHILD: Tadah. MOTHER: Tadah, you want me to read it to you? MOTHER: Cmon. MOTHER: Let's go see what it's about. MOTHER: You know what it's called? CHILD: What? MOTHER: The great cat chase. MOTHER: You know what that must be about? CHILD: Uhhuh. MOTHER: Chasing ___ cat.
3	mot	25	a	MOTHER: Isn't it? CHILD: Yeah. MOTHER: Yeah. CHILD: Now. CHILD: This goes like that. MOTHER: Mhm that's a CHILD: Now ... MOTHER: You know what this is called? CHILD: What? MOTHER: This is called ___ jack.

3	mot	26	a	MOTHER: If you got in Rachel's carriage you'd look like that because you'd be too what? CHILD: Big. MOTHER: Too big for it. MOTHER: That's right. MOTHER: And look at the little boy. MOTHER: And the kitty. MOTHER: And the lady's home. MOTHER: I wonder if that's the lady's little boy. MOTHER: I don't know. MOTHER: And now she's coming out pretending she's ___ nurse.
3	mot	27	a	MOTHER: It's called the Great cat chase. CHILD: Mhm. MOTHER: Can you find the kitty cat here? CHILD: Yeah. MOTHER: Where? MOTHER: Which one? CHILD: What's that? MOTHER: Oh, it's a camera. CHILD: What's that? MOTHER: The man's taking ___ picture, but you don't hafta worry.

3	mot	28	a	MOTHER: They want you to talk. MOTHER: Look it, do you think she's taking the baby for a walk? CHILD: Huh? MOTHER: What is she doing? MOTHER: What is she doing? CHILD: Yeah, I'll read it to you. MOTHER: You're gonna read it. MOTHER: Oh, okay, you read it to me. CHILD: What's that? MOTHER: That's just ... __ piano I think.
3	mot	29	a	MOTHER: And she's pretending to be a mommy. MOTHER: And who did she talk into being a baby? CHILD: Kitty. MOTHER: The kitty cat's being the baby. MOTHER: Does that kitty cat look happy about it? CHILD: Hunhunh. MOTHER: No. MOTHER: I don't think he wants to be a baby. MOTHER: Okay. MOTHER: So she turned around to pick __ flower, and what'd the kitty cat do?

3	mot	30	a	MOTHER: You do? MOTHER: Okay. MOTHER: Why do you hafta close up first? CHILD: Hey what's this for? CHILD: What? MOTHER: I don't know what that one's for. CHILD: Hmm. MOTHER: Hmm I wonder if you were ... MOTHER: Maybe that's for where you get money back I don't know how that works, Charlie. MOTHER: I bet you have __ tape.
3	mot	31	a	MOTHER: And she said ... MOTHER: Oh dear. MOTHER: Where has my baby gone? MOTHER: And there's her baby. MOTHER: Here's the pussy cat right behind a tree. MOTHER: But Mathilda doesn't know that. MOTHER: And so she goes running down the street. MOTHER: Help. MOTHER: Help. MOTHER: And there she sees her friend Charlie pre- tending he's __ policeman.

3	mot	32	a	MOTHER: Do you think so? CHILD: Could screw this. MOTHER: Yeah that's the wrong way though. CHILD: Hey there's some more toys up there! MOTHER: You want me to show you what the other one is? CHILD: Yeah. MOTHER: There's one more box, and I'll show you, okay? CHILD: What? CHILD: What? MOTHER: Here's ___ book.
3	mot	33	a	MOTHER: Okay, here, all the pieces that you take out we'll put in a pile. CHILD: Okay. MOTHER: Does that loosen it? MOTHER: Okay okay, what else would you do with that screwdriver? CHILD: Flashlight. MOTHER: The flashlights? MOTHER: The headlights. CHILD: Yeah, headlights. MOTHER: The headlights. MOTHER: This is like being ___ mechanic.

3	mot	34	a	MOTHER: You wanna sit here where the cashier sits? CHILD: Yes put this ... CHILD: Here. MOTHER: Okay, alright now you sit down. MOTHER: Okay. CHILD: Okay. MOTHER: Okay, alright, here's my shopping cart. MOTHER: Here's what else I'm buying. MOTHER: Are you gonna tell me how much money it is? MOTHER: Here's ___ box of napkins.
3	mot	35	a	MOTHER: And they come out the other direction. MOTHER: And, meanwhile, the little boy who was the cowboy caught the cat MOTHER: And here comes the kids through the pipe. MOTHER: Oh no! CHILD: What? MOTHER: Oh no! MOTHER: What did he do? MOTHER: What did the cowboy do? MOTHER: He dropped that kitty right on the policeman's head! MOTHER: And now the policeman has a ... a cat hat instead of ___ policeman hat.

3	mot	36	a	CHILD: What about the other things? MOTHER: You don't wanna do any more of this? MOTHER: Do you know what this is chuck? CHILD: What? MOTHER: Watch this. MOTHER: Here? MOTHER: We'll pretend this is a store and try that. MOTHER: What's this? CHILD: Typewriter? MOTHER: No, it's in ___ store.
3	mot	37	a	MOTHER: Try it there. MOTHER: Maybe ... MOTHER: Ooh! MOTHER: There you go! MOTHER: Oh Charlie. MOTHER: A lot of money. CHILD: Could put this. MOTHER: You know what we could do? CHILD: What? MOTHER: We could pretend that this was ___ store, right?

3	mot	38	a	CHILD: Why is ... why he's hiding behind a tree? MOTHER: It's hiding behind a tree, isn't it? CHILD: Why? MOTHER: I guess he doesn't want her to see her. MOTHER: What do you think? MOTHER: Look. MOTHER: She's running to tell somebody. MOTHER: And who does she tell? CHILD: Policeman. MOTHER: Policeman, and another kid on __ bike too.
3	mot	39	a	MOTHER: You use to eat that when you were little. MOTHER: There nothing in it, though. CHILD: Could I take this? MOTHER: Uhhuh. CHILD: What is it called? MOTHER: Well. MOTHER: Do ya remember? MOTHER: It's a ... CHILD: What is it? MOTHER: It's __ toy.

3	mot	40	a	MOTHER: Cause this one's very different. MOTHER: It's ... it's like yours. MOTHER: Do you remember how to do yours? CHILD: What? MOTHER: It's got all the tools. MOTHER: It comes apart. MOTHER: That's the wrench, right? MOTHER: This is the screwdriver. MOTHER: I don't know what that is. MOTHER: That's __ screwdriver.
---	-----	----	---	---

D.6 Mother determiner uses in 4-year-old corpora

age	speaker	item	ans	context
-----	---------	------	-----	---------

4	mot	1	the	Child: It's... it's a metal. Child: It's a plastic screw. Mother: It's a plastic screw. Child: It's not a real screw. Mother: Yeah. Mother: That's true. Child: Where does... where does this go? Mother: I think that goes from the front but ... Mother: That's the ... windshield I think. Mother: Do you wanna finish taking that screw out so you can take __ side out?
4	mot	2	the	Mother: Is that very nice? Mother: To hit kittys? Mother: I sure wouldn't like that. Child: Mhm. Mother: No. Mother: It is not really nice to hit animals. Mother: People don't do that. Mother: Only mean people who do that. Mother: So. Mother: Here's __ kitty.

4	mot	3	the	Mother: Oh alright open it up now. Mother: Let's go. Mother: Okay can I give you a little help? Mother: Okay now put the other tire on. Mother: Go ahead. Mother: Put it on. Mother: Good, you can use a screw driver or you know you can even use your wrist, you can even do it like this for a while until you can turn it like that. Mother: That's it. Mother: Good. Mother: And then you tighten it with ___ screw driver, I think that's a better idea.
4	mot	4	the	Child: Three. Mother: Three wheels. Mother: What happens next? Child: What? Mother: What's the policeman saying to the kitty? Child: I don't know. Mother: Come here, kitty, kitty, kitty, kitty, kitty, kitty. Mother: He's gonna try to catch it. Mother: What happens? Mother: ___ kitty what?

4	mot	5	the	Child: Right here. Mother: Okay come over..... over here. Mother: Are you gonna be the check out lady? Child: Yeah. Mother: Okay, alright, should I get some things? Child: Yeah. Mother: And should I... Mother: What... what do you think we oughta buy Helen? Child: I don't know. Mother: Well, let's look at what's in ___ store over here.
4	mot	6	the	Child: He's in a tree. Child: And who needs help? Mother: Who needs help? Child: I don't know. Mother: Who's this? Child: The policeman. Mother: Yeah. Child: Why he needs help? Mother: Well, he looks like he's about to fall out of the tree. Mother: He's trying to get ___ cat.

4	mot	7	the	Mother: Let's take the car over here where they can see. Child: Hey, look at that. Mother: Yeah, that's the same one that we have. Child: Yeah. Child: Are we gonna put it together? Mother: Okay. Child: But it broke. Mother: Well remember... what we do at home with it? Mother: That's it. Mother: You found ___ screwdriver.
4	mot	8	the	Mother: Okay, here you wanna take a look at this? Child: Mhm. Mother: Okay. Child: What is the ... Mother: Come here. Child: Of the, one thing. Mother: This is the other one thing. Mother: The store. Mother: Come. Mother: You wanna read ___ story with me?

4	mot	9	the	Mother: No. Mother: The cat ... Mother: Well I don't know I guess that the cat jumped on his... his head. Child: (laughs). Mother: (laughs) And then... and then he got away and he's going up the tree again. Mother: Oh, is he upset, huh? Mother: Oh look at this, the policeman climbed all the way up the tree. Mother: And there.. hangs the cat. Mother: And then he got it. Mother: That's ___ way you're sposta catch a cat.
---	-----	---	-----	---

4	mot	10	the	Mother: These are like you know the ones that we have the tiny boxes and cans of food? Mother: I'm taking some money to play store, okay? Child: Okay. Mother: I'll take a number one, a one dollar bill... and a two. Child: You. Mother: Oh, you know what? Mother: I have an idea, John. Mother: Why don't we use this as the counter to pay? Mother: Because they can't see you when you are behind there. Mother: ___ camera can't see you.
---	-----	----	-----	--

4	mot	11	the	Mother: He can't even walk. Mother: What's that that he has? Mother: A crutch . Child: What is these ... Mother: Now he's walking with a crutch. Child: That's not a crutch that's a branch. Mother: Yeah. Mother: But he's using it for a crutch. Child: Yeah. Mother: Because he can't walk cause he fell off a tree and, you know, got all wet in ___ puddle.
4	mot	12	the	Mother: Why don't you make coffee for what? Mother: Oh probably about.. ten cents? Mother: They'll come up when you push the next button dear. Child: Alright. Mother: Ten cents for.. cereal? Child: Ten. Mother: How bout that? Mother: And how much are you going to charge me for the cocoa? Child: Oh Mom, look at the box. Mother: Alright let's see what ___ box says.

4	mot	13	the	Child: Alright. Mother: No. Mother: It won't stay up. Mother: No. Mother: It won't stand up. Mother: Go ahead. Child: No, that's okay... that's okay. Child: You can just hold it. Mother: Okay . Mother: I'll be ___ gas station.
---	-----	----	-----	--

4	mot	14	the	Mother: Are you a youngster? Child: Yeah. Mother: No? Mother: Youngsters will find out in this delightful picture story without words. Mother: The exasperated wily cat, deter... determined little girl, and this Keystone copy policeman come to her rescue all come hilariously to life in Mercer Mayer's exuberant drawings. Mother: Ah.. what're they doing there? Child: I don't know. Mother: ... What is this? Child: Cat. Mother: And what's ___ cat in?
---	-----	----	-----	---

4	mot	15	the	Mother: Oh she turns to pick the flowers and what's... look what happens Mother: What happens? Mother: Can you tell ... Mother: You... you've seen the book before. Mother: Why don't you tell me the story? Child: Okay. Child: The cat runs away and hides. Mother: Mhm. Mother: But where does it hide? Mother: Behind __ tree?
4	mot	16	the	Mother: Come here kitty kitty kitty. Mother: Look what the cat is going into. Child: A rain barrel. Mother: A rain barrel it looks like. Mother: Yeah. Mother: Like a big pipe. Mother: The cat looks pretty angry. Mother: ... Oh. Mother: Somebody's crawling in. Mother: __ police woman is crawling in.

4	mot	17	the	Mother: Thank you very much. Mother: I'm gonna put it back cause we don't own this money. Mother: Okay. Mother: Do you wanna go play with some of the other toys? Child: No. Child: No. Child: You hafta bring this home. Mother: Oh. Mother: Put it in a bag for me. Mother: You gonna put it in a bag just like you do in ___ grocery store?
---	-----	----	-----	--

4	mot	18	the	Mother: And the little boy that's got the Batman shirt on is very angry. Mother: And so he stuck his tongue out. Mother: And then.. what happened? Mother: The cat fell. Mother: And the little boy fell. Mother: See how the branch broke? Mother: That's what could happen when you go climbing on trees and don't go on strong l... limbs, the branch breaks and goes.... Mother: Crash. Mother: And the cat fell. Mother: And the little girl caught him, and ___ policeman fell and landed right on the little boy with the Batman shirt.
---	-----	----	-----	--

4	mot	19	the	Child: He... he ... he ... he's telling ... he's telling that the cat was in the stroller and was ... Mother: Oh. Mother: Oh. Mother: Is he gonna help her find the cat? Child: Yeah. Mother: Oh. Mother: I see. Child: They're right here. Mother: Oh. Mother: But she doesn't see ___ cat, does she?
4	mot	20	the	Mother: They're all ... Mother: Look at that. Mother: The cat is swinging on the branch, and the policewoman is grabbing for her. Mother: Oh no! Mother: What shall we do? Mother: But she got her. Mother: Oh, I think this is sposta be a boy. Mother: This one's a boy. Mother: But I don't know. Mother: She grabbed ___ cat.

4	mot	21	a	Mother: And the policeman's walking with a cane. Mother: He's so exasperated. Mother: He's been through so much, trying to catch the cat. Mother: And the cat's walking along behind like a little prince, huh? Child: Yeah. Mother: Yeah. Child: Turn the page. Mother: And then there's the boy in the carriage. Mother: And they all go home.. to the little girl's house. Mother: And now she brings out ___ pitcher of.. what do ya think?
4	mot	22	a	Mother: Mhm. Child: I like these. Child: Coffee? Mother: You want coffee? Mother: Okay, let me add that up. Mother: You want coffee and the cookies? Mother: Anything else ? Child: Cheese . Child: What is this? Mother: That's... ___ rabbit drink.

4	mot	23	a	Mother: Something not very nice. Mother: He's sticking his tongue out. Child: To the policeman ? Mother: No. Mother: That's just a little boy. Mother: That's not really a policeman, Davey. Mother: It's just a little boy dressed up as a policeman. Mother: She the great big funny shoes and the big.. blue shirt. Mother: That's not really a policeman. Mother: He just has ___ police outfit on.
---	-----	----	---	--

4	mot	24	a	Child: Yeah this one has a flat tire. Mother: Yeah. Child: We gonna put this on. Mother: Oh alright open it up now. Mother: Let's go. Mother: Okay can I give you a little help? Mother: Okay now put the other tire on. Mother: Go ahead. Mother: Put it on. Mother: Good, you can use ___ screw driver or you know you can even use your wrist, you can even do it like this for a while until you can turn it like that.
4	mot	25	a	Mother: Which are the quarters? Mother: Do you know? Child: Yes. Mother: What's that one? Mother: What is it? Child: Quarter. Mother: Right where's a dime? Child: Dime. Mother: Right. Mother: Where's ___ penny ?

4	mot	26	a	Mother: Yeah? Mother: You know something? Mother: I think this car comes apart. Mother: You can use the tools. Mother: Oh, look at this. Child: What? Mother: The windshield came off. Child: What what is this? Mother: Well let's see. Mother: Is that ___ tool you can use?
4	mot	27	a	Child: You do it, please. Mother: See ? Mother: See, and in fact if you use two hands, it might be a little bit easier for you. Mother: Should ... Mother: Let's try to put another wheel in. Mother: Okay ? Mother: Alright. Child: I'm gonna put it in. Mother: Okay, alright. Mother: I'll hold the car for you and you get ___ screw, that's right.

4	mot	28	a	Mother: Hurray: . Child: Oop . Child: Hurray now the next wheel . Mother: Okay. Mother: Next wheel. Mother: Inside here is another game that they said afterwards we can have a chance to play. Child: Yeah. Mother: It's playing store I think with money. Child: Yeah. Mother: Yeah, we have a chance to do three things in this room, playing with the car, playing store, and reading __ book.
4	mot	29	a	Mother: Try again. Child: That's because you moved it there. Mother: I moved it ? Child: Yeah. Mother: Xavia. Mother: Pick it up and let me show you something. Mother: Lift up the motor a minute. Child: No. Mother: No. Mother: .. You want __ hint, let me know.

4	mot	30	a	Child: Tight yet? Mother: No, it's not. Mother: Now the... the... the red ones go behind right? Mother: Now let's take the black one. Mother: Where's the black screw? Mother: Do you know? Mother: We need another black screw. Mother: Or did we.. do we just have one black screw? Child: Here's a black screw. Mother: No that's ___ red screw.
4	mot	31	a	Mother: Oh, did you do this with Daddy last week? Child: Yeah. Mother: Oh. Mother: That's really nice. Child: Oh. Child: That's the wrong way. Mother: There you go. Mother: Hmm. Mother: Oh. Mother: They even have ___ trunk with a spare tire in it.

4	mot	32	a	Mother: A what? Child: Backer car. Mother: A backer car? Mother: What's that? Child: I need ... Child: Let's see ... Child: I this... I need this. Mother: Oh. Mother: It's so small! Mother: That's called ___ jack.
4	mot	33	a	Mother: Want me to do this one? Mother: Or do you wanna do it? Child: Do this one. Child: And I'll take ... Child: Let me take the screws off. Mother: You know what this part of the car's called? Child: It's only a screw. Mother: Well. Mother: It's a screw on this car. Mother: But on real cars it's called ___ hubcap.

4	mot	34	a	Mother: I'll come and buy something from you . Child: No . Mother: I'm gonna come and buy something. Child: I don't wanna. Mother: Okay do you wanna buy something from me? Child: No. Mother: Huh? Child: What? Mother: Would you like to buy something from me? Mother: I'll be ___ store lady, okay?
4	mot	35	a	Mother: You spent a dollar and fifty cents. Mother: And you have two dollars, right? Mother: Okay I hafta give you change. Mother: A dollar fifty and twenty five and and another twenty five makes another fifty cents and that makes two dollars. Mother: And I take the two dollars and you have the change. Mother: But you have the grocerys to go with it, right? Child: I wanna carry these home in my hands. Mother: Oh you're gonna carry them home in your hands? Mother: That's gonna be hard. Mother: Wouldn't you like me to give you ___ bag?

4	mot	36	a	Mother: Okay, Child: Here's two quarters. Mother: Oh thank you. Mother: Thank you. Child: Now I gotta put it back in the cash register . Mother: Okay . Mother: Okay. Child: What are those? Mother: I don't know. Mother: Should we... should we put the packages in ____ bag now?
4	mot	37	a	Child: I don't know how much money. Mother: Well. Mother: Why don't you ring it up on the cash register? Mother: How much is this.., sir? Child: Umm.. twelve. Mother: Okay. Mother: Ring that... ring that up. Mother: You hafta ring up each thing. Child: Okay. Mother: Have you ever run ____ grocery store before?

4	mot	38	a	Mother: That ... that's so where the mother... the mother cat.. bites the kittens, always on the back of the neck. Mother: She carries them around that way. Child: Yeah and she says she liked them. Mother: Mhm. Child: He's gonna fall . Mother: Wow , and there the branch of the tree is breaking off. Mother: The policeman's going to fall. Mother: He did fall. Mother: But the cat landed in the arms of the little girl. Mother: There's ___ shoe in the air.
---	-----	----	---	--

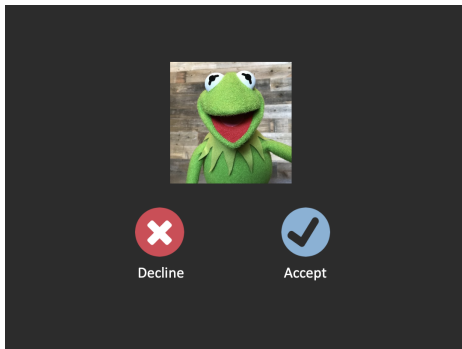
4	mot	39	a	Mother: Hmm? Child: What's this one? Mother: That's a wrench. Mother: That. Mother: You could have used that for some of them but Mother: Some of the screws did ... Mother: But instead you used the screw driver. Mother: Okay. Mother: Let's figure out. Mother: Does this look like ___ car?
4	mot	40	a	Mother: Alright. Mother: It's called The great cat chase. Mother: A wordless book by Mercer Mayer. Mother: What kind of a name is that? Child: Mercer Mayer. Mother: What's that? Child: Kitty. Mother: Yeah. Mother: Okay. Mother: This little girl is ... is rolling her kitty in ___ baby carriage.

Appendix E: Stimuli for Study 6

This appendix lays out the production procedure for a participant.

All scenes of Kermit talking are lip-synchronized to create a more real communicative context. Furthermore, a professional voice actor was invited to dub the scenes to enhance the effect and guarantee a more natural, social setup.¹

E.1 Meeting Kermit



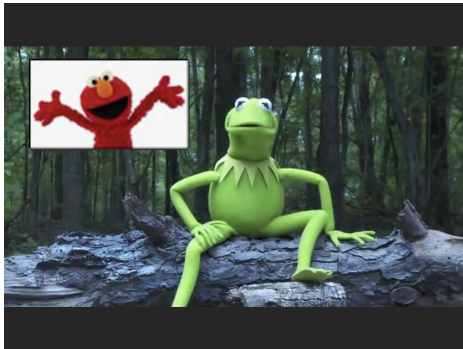
(Phone ringing)

¹Based on reactions and feedback from child participants, they indeed seemed to believe that the figure they were talking to was Kermit the frog.



Hey there! I'm Kermit the frog!

E.2 Introducing the setup



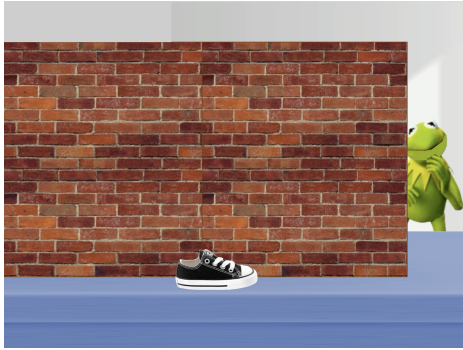
This is my friend Elmo. We always play together. Now I need some help. We're going to see some places with walls. Elmo said there were interesting things behind these walls, but he didn't tell me what was there. So I really need your help. When you see something that

I can't see, tell me what it is.

E.3 Practice trials

There are three practice trials in total. The first two trials involve simple questions about the identity of a single object. The last trial includes both an object identity question and an event-related question.

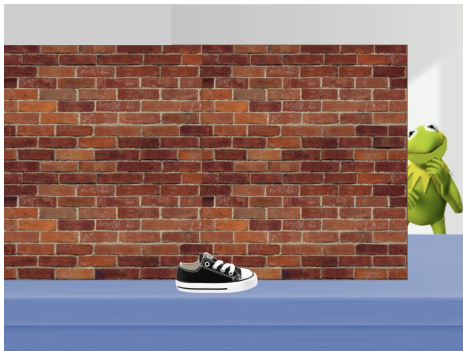
During the practice trials, the experimenter would only chime in when child participants were stuck, either due to their unfamiliarity with the situation or the character or because of their shyness in communicating with Kermit. Moreover, the experimenter provided no feedback regarding their responses.



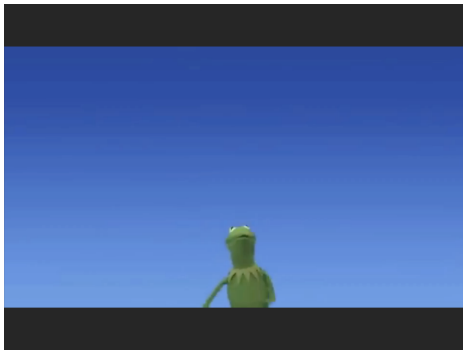
Practice trial type1 (x2):

Here we are. What a lovely place! I really want to know what's behind this wall, but I can't see it from here. What do you see?

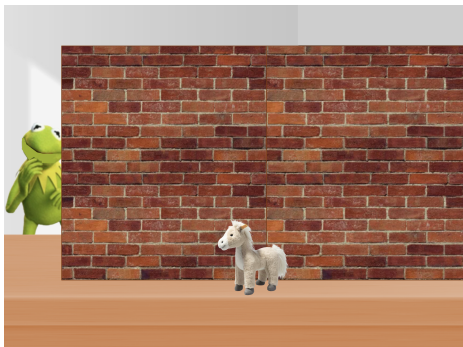
(Expected response: A shoe.)



Okay. Thanks!



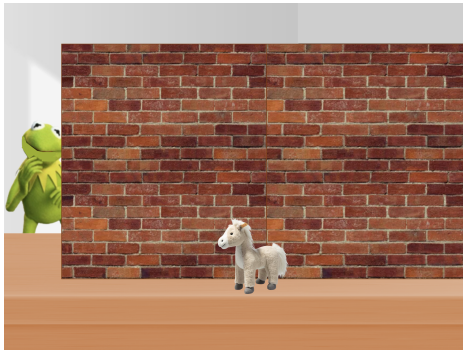
Alright. Let's go see the next spot.



Practice trial type2 (x1): Here we are.

What a nice place! I really want to know what's behind this wall, but I can't see it from here. What do you see?

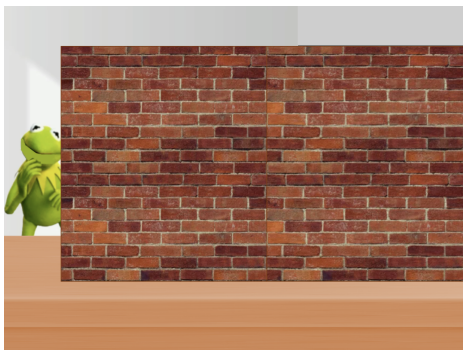
(Expected response: A horse/pony.)



Okay.

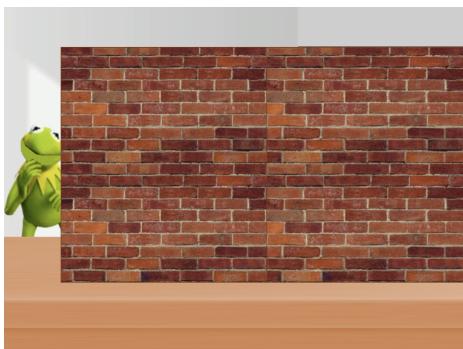


(Elmo takes the object, laughing.)



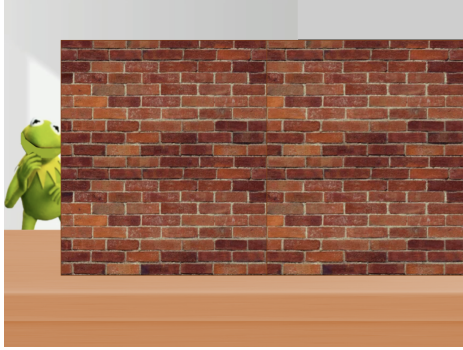
Wait. Did you hear something? It sounded like someone's behind the wall. Who was it?

(Expected response: Elmo).



What happened?

(Expected response: Elmo took **the** horse).



Okay. Thanks!

E.4 Test trials

Before the test trials started, the experimenter would explain to the participant that he had work to do and would not join in the conversation. Crucially, he always turned sideways to prevent the participant from being distracted by his presence. This is to guarantee that during the test trials, the participant would solely engage with Kermit and consider Kermit's perspective.

There are twelve test trials in total. While all test trials include both an object identity question and an event-related question, half of them feature two unique objects and the other half feature two non-unique or identical objects.

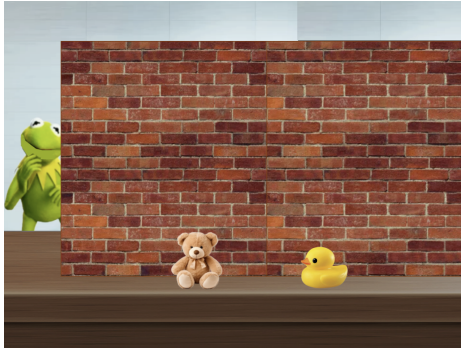
Below is a list presenting the twelve referential contexts.

Unique referent contexts (the object taken away in bold)	Non-unique referent contexts (one of the two objects taken away)
a bottle and a mug	two bags
a ball and a car	two boxes
an orange and an apple	two bananas
a bear and a duck	two books
a carrot and a hat	two boats
a truck and a cake	two bowls

Table E.4: List of referential contexts in Study 6

The order of these test trials are pseudorandomized, such that participants cannot anticipate what an upcoming referential context will be. One of two pseudorandomized trial orders is randomly assigned to each participant.

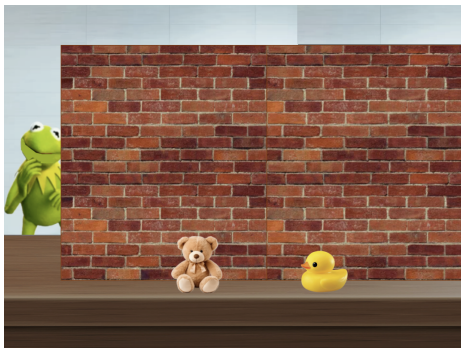
Here, I provide examples for each of the two test trial types.



Test trial type1:

Here we are. What a nice place! I really want to know what's behind this wall, but I can't see it from here. What do you see?

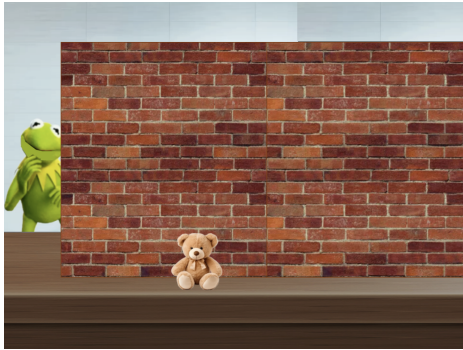
(Expected response: A teddy bear and a rubber duckie.)



Okay.



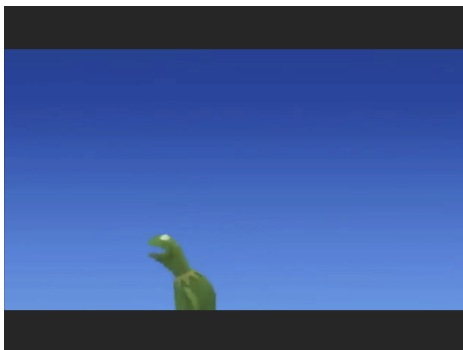
(Elmo takes one of the objects, laughing.)



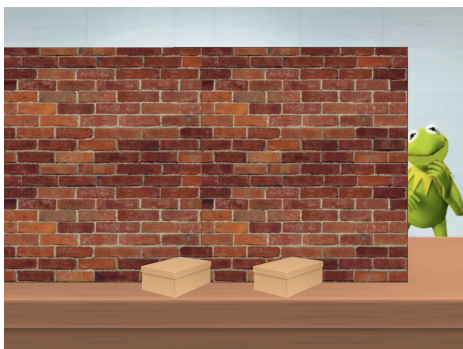
Did you hear something? It sounded like Elmo's behind the wall. What happened?
(Expected response: Elmo took **the** rubber duckie).



Okay. Thanks!

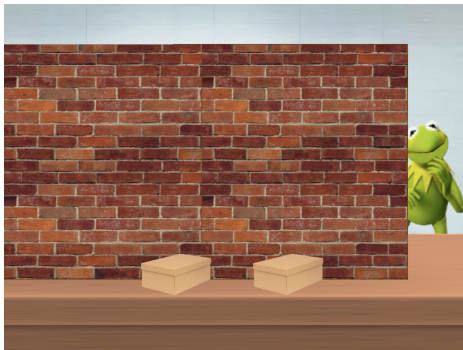


Alright. Let's go see the next spot.



Test trial type2:

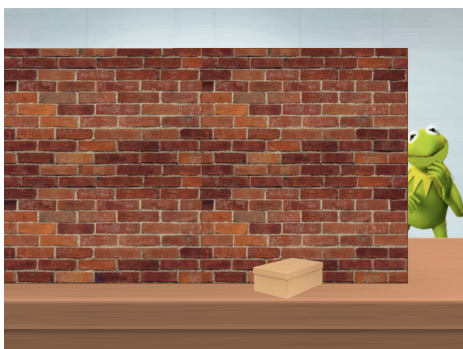
Here we are. What a lovely place! I really want to know what's behind this wall, but I can't see it from here. What do you see?
(Expected response: {Boxes/Two boxes}.)



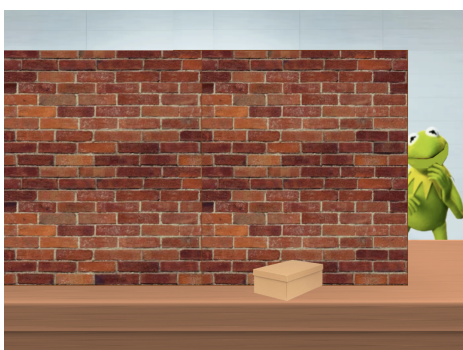
Okay.



(Elmo takes one of the objects, laughing.)

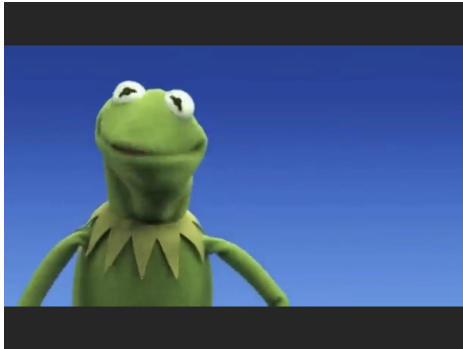


Did you hear something? It sounded like
Elmo's behind the wall. What happened?
(Expected response: Elmo took {**a box/one
of the boxes**}).

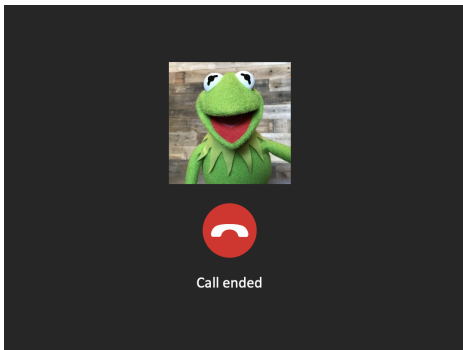


Okay. Thanks!

E.5 Thanks and goodbye



Thank you so much for telling me about the things I couldn't see. I'll have to go now. Bye!



(Phone hung up)

Bibliography

- Abbott, B. (2006). Where have some of the presuppositions gone. *Drawing the boundaries of meaning: Neo-Gricean studies in pragmatics and semantics in honor of Laurence R. Horn*, 80, 1–20.
- Abusch, D. (2002). Lexical alternatives as a source of pragmatic presuppositions. *Semantics and linguistic theory*, 1–19.
- Adibpour, P., & Hochmann, J.-R. (2023). Infants’ understanding of the causal power of agents and tools. *Proceedings of the National Academy of Sciences*, 120(50), e2309669120.
- Ahn, D., & Arunachalam, S. (2020). Anaphoric ‘that’: Difference between adults and children. *Proceedings of the Boston University Conference on Language Development (BUCLD)*, 44, 28–39.
- Aravind, A. (2018). *Presuppositions in context* [Doctoral dissertation, Massachusetts Institute of Technology].
- Atkinson, E., Wagers, M. W., Lidz, J., Phillips, C., & Omaki, A. (2018). Developing incrementality in filler-gap dependency processing. *Cognition*, 179, 132–149.
- Babineau, M., & Christophe, A. (2022). Preverbal infants’ sensitivity to grammatical dependencies. *Infancy*, 27(4), 648–662.
- Bassano, D., Maillochon, I., Korecky-Kröll, K., Van Dijk, M., Laaha, S., Dressler, W. U., & Van Geert, P. (2011). A comparative and dynamic approach to the

- development of determiner use in three children acquiring different languages. *First Language*, 31(3), 253–279.
- Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, 81(2), B33–B44.
- Brezack, N., Meyer, M., & Woodward, A. L. (2021). Three-year-olds’ perspective-taking in social interactions: Relations with socio-cognitive skills. *Journal of Cognition and Development*, 22(4), 537–560.
- Brodbeck, C., Gwilliams, L., & Pykkänen, L. (2015). Eeg can track the time course of successful reference resolution in small visual worlds. *Frontiers in psychology*, 6, 1787.
- Brody, G., & Csibra, G. (2025). Discourse referents in infancy. *Psychological Review*.
- Brown, R. (1973). *A first language: The early stages*. Harvard University Press.
- Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic bulletin & review*, 16, 893–900.
- Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, 112(2), 337–342.
- Campbell, A. L., Brooks, P., & Tomasello, M. (2000). Factors affecting young children’s use of pronouns as referring expressions. *Journal of Speech, Language, and Hearing Research*, 43(6), 1337–1349.
- Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.
- Carlson, S. M., Claxton, L. J., & Moses, L. J. (2015). The relation between executive function and theory of mind is more than skin deep. *Journal of Cognition and Development*, 16(1), 186–197.
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children’s theory of mind. *Child development*, 72(4), 1032–1053.

- Cheng, L. L.-S., & Sybesma, R. (1999). Bare and not-so-bare nouns and the structure of np. *Linguistic inquiry*, 30(4), 509–542.
- Chierchia, G., Guasti, M. T., Gualmini, A., Meroni, L., Crain, S., & Foppolo, F. (2004). Semantic and pragmatic competence in children’s and adults’ comprehension of or. In *Experimental pragmatics* (pp. 283–300). Springer.
- Choi, Y.-j., Song, H.-j., & Luo, Y. (2018). Infants’ understanding of the definite/indefinite article in a third-party communicative situation. *Cognition*, 175, 69–76.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
- Clark, H. H. (1977). Bridging. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking: Readings in cognitive science* (pp. 411–420). Cambridge University Press.
- Cohn, N. (2020). Visual narrative comprehension: Universal or not? *Psychonomic Bulletin & Review*, 27(2), 266–285.
- Crain, S., McKee, C., & Emiliani, M. (1990). Visiting relatives in Italy. In *Language processing and language acquisition* (pp. 335–356). Springer.
- Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & language*, 25(2), 141–168.
- Dayal, V., & Jiang, L. J. (2022). The puzzle of anaphoric bare nouns in Mandarin: A counterpoint to index! *Linguistic inquiry*, 54(1), 147–167.
- de Cat, C. (2013). Egocentric definiteness errors and perspective evaluation in preschool children. *Journal of Pragmatics*, 56, 58–69.
- De Cat, C. (2015). The cognitive underpinnings of referential abilities. In *The acquisition of reference* (pp. 263–283). John Benjamins.
- Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, 44(1-2), 1–42.
- DeLoache, J. S., & Burns, N. M. (1994). Early understanding of the representational function of pictures. *Cognition*, 52(2), 83–110.

- Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition*, 11(2), 159–184.
- Dieuleveut, A., Van Dooren, A., Cournane, A., & Hacquard, V. (2022). Finding the force: How children discern possibility and necessity modals. *Natural Language Semantics*, 30(3), 269–310.
- Dink, J. W., & Ferguson, B. (2015). Eyetrackingr: An r library for eye-tracking data analysis.
- Dowty, D. (2003). *The dual analysis of adjuncts/complements in categorical grammar* (E. Lang, C. Maienborn, & C. Fabricius-Hansen, Eds.). De Gruyter Mouton.
- Dudley, R. (2017). *The role of input in discovering presupposition triggers: Figuring out what everybody already knew* [Doctoral dissertation, University of Maryland, College Park].
- Dye, C., Kedar, Y., & Lust, B. (2019). From lexical to functional categories: New foundations for the study of language development. *First Language*, 39(1), 9–32.
- Emslie, H. C., & Stevenson, R. J. (1981). Pre-school children’s use of the articles in definite and indefinite referring expressions. *Journal of Child language*, 8(2), 313–328.
- Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of personality and social psychology*, 87(3), 327.
- Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, 40, 760–768.

- Fenson, L., Bates, E., Dale, P. S., Hartung, J. P., Pethick, S. J., Reilly, J. S., & Thal, D. J. (1993). *The macarthur communicative development inventories*. Singular Publishing Group.
- Fillmore, C. J. (1968). The case for case. In E. W. Bach & R. T. Harms (Eds.), *Universals in linguistic theory* (pp. 1–88). New York: Holt, Rinehart; Winston.
- Fisher, C., Gertner, Y., Scott, R. M., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(2), 143–149.
- Fisher, C., & Tokura, H. (1995). The given-new contract in speech to infants. *Journal of Memory and Language*, 34(3), 287–310.
- Fraurud, K. (1990). Definiteness and the processing of noun phrases in natural discourse. *Journal of semantics*, 7(4), 395–433.
- Freeman, N. H., Lewis, C., & Doherty, M. J. (1991). Preschoolers’ grasp of a desire for knowledge in false-belief prediction: Practical intelligence and verbal report. *British Journal of Developmental Psychology*, 9(1), 139–157.
- Friederici, A. D., Mueller, J. L., & Oberecker, R. (2011). Precursors to natural grammar learning: Preliminary evidence from 4-month-old infants. *PLoS One*, 6(3), e17920.
- Fukumura, K., & van Gompel, R. P. (2012). Producing pronouns and definite noun phrases: Do speakers use the addressee’s discourse model? *Cognitive Science*, 36(7), 1289–1311.
- Gagarina, N., & Bohnacker, U. (2022). A new perspective on referentiality in elicited narratives: Introduction to the special issue. *First Language*, 42(2), 171–190.
- Gagliardi, A., Mease, T. M., & Lidz, J. (2016). Discontinuous development in the acquisition of filler-gap dependencies: Evidence from 15-and 20-month-olds. *Language Acquisition*, 23(3), 234–260.

- Ganea, P. A., Allen, M. L., Butler, L., Carey, S., & DeLoache, J. S. (2009). Toddlers' referential understanding of pictures. *Journal of experimental child psychology*, 104(3), 283–295.
- Geffen, S., & Mintz, T. H. (2015). Can you believe it? 12-month-olds use word order to distinguish between declaratives and polar interrogatives. *Language Learning and Development*, 11(3), 270–284.
- Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, 73(2), 135–176.
- Gleitman, L. (1990). The structural sources of verb meanings. *Language acquisition*, 1(1), 3–55.
- Goldberg, A. E. (2002). Surface generalizations: An alternative to alternations.
- Gómez, R., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, 7(2), 183–206.
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological science*, 13(5), 431–436.
- Grigoroglou, M., & Papafragou, A. (2019). Interactive contexts increase informativeness in children's referential communication. *Developmental psychology*, 55(5), 951.
- Grosse, G., Moll, H., & Tomasello, M. (2010). 21-month-olds understand the cooperative logic of requests. *Journal of Pragmatics*, 42(12), 3377–3383.
- Guerriero, A. S., Oshima-Takane, Y., & Kuriyama, Y. (2006). The development of referential choice in english and japanese: A discourse-pragmatic perspective. *Journal of child language*, 33(4), 823–857.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 274–307.

- Gundel, J. K., & Johnson, K. (2013). Children's use of referring expressions in spontaneous discourse: Implications for theory of mind development. *Journal of Pragmatics*, 56, 43–57.
- Hamburger, H., & Crain, S. (1982). Relative acquisition. In *Language development vol. i: Syntax and semantics* (pp. 245–274). Hillsdale, NJ: Erlbaum.
- Hawkins, J. A. (1978). *Definiteness and indefiniteness*. Croom Helm, London.
- He, A. X. (2015). *Verb learning under guidance* [Doctoral dissertation, University of Maryland].
- Heim, I. R. (1982). *The semantics of definite and indefinite noun phrases* [Doctoral dissertation, University of Massachusetts Amherst].
- Henning, A., Spinath, F. M., & Aschersleben, G. (2011). The link between preschoolers' executive function and theory of mind and the role of epistemic states. *Journal of experimental child psychology*, 108(3), 513–531.
- Hirzel, M. R. (2022). *Island constraints: What is there for children to learn?* [Doctoral dissertation, University of Maryland, College Park].
- Hoff, E. (2010). Context effects on young children's language use: The influence of conversational setting and partner. *First language*, 30(3-4), 461–472.
- Hogrefe, G.-J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child development*, 567–582.
- Hohle, B., Schmitz, M., Santelmann, L. M., & Weissenborn, J. (2006). The recognition of discontinuous verbal dependencies by german 19-month-olds: Evidence for lexical and structural influences on children's early processing capacities. *Language Learning and Development*, 2(4), 277–300.
- Huang, C.-c. (2012). Referential choice and informativeness in mother-child conversation: A focus on the mother. *Language Sciences*, 34(3), 319–338.

- Huang, C.-T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic inquiry*, 531–574.
- Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental psychology*, 43(6), 1447.
- Hughes, M. E., & Allen, S. E. M. (2013). The effect of individual discourse-pragmatic features on referential choice in child english. *Journal of Pragmatics*, 56, 15–30.
- Hughes, M. E., & Allen, S. E. M. (2015). The incremental effect of discourse-pragmatic sensitivity on referential choice in the acquisition of a first language. *Lingua*, 155, 43–61.
- Hyams, N. (1996). The underspecification of functional categories in early grammar. In H. Clahsen (Ed.), *Generative perspectives on language acquisition* (pp. 91–128). John Benjamins Publishing Company.
- Ikeda, A., Kanakogi, Y., & Hirai, M. (2022). Visual perspective-taking ability in 7-and 12-month-old infants. *Plos one*, 17(2), e0263653.
- Ionin, T. (2003). *Article semantics in second language acquisition* [Doctoral dissertation, Massachusetts Institute of Technology].
- Jackendoff, R. S. (1992). *Semantic structures* (Vol. 18). MIT press.
- Jenks, P. (2018). Articulated definiteness without articles. *Linguistic inquiry*, 49(3), 501–536.
- Jin, K.-s., & Fisher, C. (2014). Early evidence for syntactic bootstrapping: 15-month-olds use sentence structure in verb learning. *BUCLD 38 Online Proceedings Supplement*.
- Karmiloff-Smith, A. (1979). *A functional approach to child language: A study of determiners and reference*. Cambridge University Press.

- Kingsbury, P., Strassel, S., McLemore, C., & McIntyre, R. (1997). Callhome American English transcripts, LDC97T14. *Philadelphia: Linguistic Data Consortium*.
- Koenig, J.-P., Mauner, G., & Bienvenue, B. (2003). Arguments for adjuncts. *Cognition*, 89(2), 67–103.
- Lempers, J. D., Flavell, E. R., & Flavell, J. H. (1977). The development in very young children of tacit knowledge concerning visual perception. *Genetic Psychology Monographs*, 95(1), 3–53.
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago press.
- Levin, B., & Hovav, M. R. (2005). *Argument realization* (Vol. 10). Citeseer.
- Levin, B., & Rappaport, M. (1988). Nonevent-er nominals: A probe into argument structure. *Linguistics*, 26(6), 1067–1084.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge university press.
- Lewis, D. (1979). Scorekeeping in a language game. *Journal of philosophical logic*, 8, 339–359.
- Lewis, S., Hacquard, V., & Lidz, J. (2017). “think” pragmatically: Children’s interpretation of belief reports. *Language Learning and Development*, 13(4), 395–417.
- Liberman, Z., Woodward, A. L., Keysar, B., & Kinzler, K. D. (2017). Exposure to multiple languages enhances communication skills in infancy. *Developmental science*, 20(1), e12420.
- Lidz, J., & Gagliardi, A. (2015). How nature meets nurture: Universal grammar and statistical learning. *Annu. Rev. Linguist.*, 1(1), 333–353.
- Lidz, J., White, A. S., & Baier, R. (2017). The role of incremental parsing in syntactically conditioned word learning. *Cognitive Psychology*, 97, 62–78.

- Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, 108(3), 732–739.
- Luo, Y., & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition*, 105(3), 489–512.
- Marantz, A. (1981). *On the nature of grammatical relations* [Doctoral dissertation, Massachusetts Institute of Technology].
- Maratsos, M. P. (1974). Preschool children's use of definite and indefinite articles. *Child development*, 45(2), 446–455.
- Maratsos, M. P. (1976). *The use of definite and indefinite reference in young children: An experimental study of semantic acquisition*. Cambridge University Press.
- Matthews, D., Butcher, J., Lieven, E., & Tomasello, M. (2012). Two-and four-year-olds learn to adapt referring expressions to context: Effects of distracters and feedback on referential communication. *Topics in cognitive science*, 4(2), 184–210.
- Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2006). The effect of perceptual availability and prior discourse on young children's use of referring expressions. *Applied Psycholinguistics*, 27(3), 403–422.
- 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.
- Matthewson, L., Bryant, T., & Roeper, T. (2001). A salish stage in the acquisition of English determiners: Unfamiliar 'definites'. *University of Massachusetts Occasional Papers in Linguistics*, 27(1), 9.
- Merchant, J. (1999). *The syntax of silence: Sluicing, islands, and identity in ellipsis*. University of California, Santa Cruz.

- Merchant, J. (2000). Islands and lf-movement in greek sluicing. *Journal of Greek Linguistics*, 1(1), 41–64.
- Mintz, T. H. (2006). Finding the verbs: Distributional cues to categories available to young learners. *Action meets word: How children learn verbs*, 31, 63.
- Mishina-Mori, S. (2007). Argument representation in japanese/english simultaneous bilinguals: Is there a crosslinguistic influence. *Proceedings of the 31st annual Boston University conference on language development*, 441–450.
- Modyanova, N., & Wexler, K. (2007). Semantic and pragmatic language development: Children know ‘that’ better. *Proceedings of the 2nd conference on generative approaches to language acquisition–north america (galana 2)*, 297–308.
- Moll, H., Carpenter, M., & Tomasello, M. (2014). Two-and 3-year-olds know what others have and have not heard. *Journal of Cognition and Development*, 15(1), 12–21.
- Moll, H., & Meltzoff, A. N. (2011). How does it look? level 2 perspective-taking at 36 months of age. *Child development*, 82(2), 661–673.
- Moll, H., Richter, N., Carpenter, M., & Tomasello, M. (2008). Fourteen-month-olds know what “we” have shared in a special way. *Infancy*, 13(1), 90–101.
- Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, 24(3), 603–613.
- Moyer, M., Harrigan, K., Hacquard, V., & Lidz, J. (2015). 2-year-olds’ comprehension of personal pronouns. *Online supplemental proceedings of the 39th Boston university conference on language development*.
- Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of child language*, 17(2), 357–374.
- Neale, S. (2004). This, that, and the other. In M. Reimer & A. Bezuidenhout (Eds.), *Descriptions and beyond* (pp. 68–188). Oxford University Press.

- Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive science*, 14(1), 11–28.
- Nilsen, E. S., & Graham, S. A. (2009). The relations between children’s communicative perspective-taking and executive functioning. *Cognitive psychology*, 58(2), 220–249.
- Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, 78(2), 165–188.
- Omaki, A., & Lidz, J. (2015). Linking parser development to acquisition of syntactic knowledge. *Language Acquisition*, 22(2), 158–192.
- Omaki, A., Orita, N., & Lidz, J. (2015). *Beyond artificial language: Statistical learning of an english non-adjacent syntactic dependency* [Unpublished manuscript].
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *science*, 308(5719), 255–258.
- Orvig, A. S., Marcos, H., Morgenstern, A., Hassan, R., Leber-Marin, J., & Parès, J. (2010). Dialogical beginnings of anaphora: The use of third person pronouns before the age of 3. *Journal of pragmatics*, 42(7), 1842–1865.
- Papafragou, A. (2006). From scalar semantics to implicature: Children’s interpretation of aspectuals. *Journal of child language*, 33(4), 721–757.
- Papeo, L., Vettori, S., Serraille, E., Odin, C., Rostami, F., & Hochmann, J.-R. (2024). Abstract thematic roles in infants’ representation of social events. *Current Biology*, 34(18), 4294–4300.
- Perkins, L. (2019). *How grammars grow: Argument structure and the acquisition of non-basic syntax* [Doctoral dissertation, University of Maryland, College Park].
- Perkins, L., Feldman, N. H., & Lidz, J. (2022). The power of ignoring: Filtering input for argument structure acquisition. *Cognitive Science*, 46(1), e13080.

- Perkins, L., Knowlton, T., Williams, A., & Lidz, J. (2024). Thematic content, not number matching, drives syntactic bootstrapping. *Language Learning and Development*, 1–31.
- Perkins, L., & Lidz, J. (2020). Filler-gap dependency comprehension at 15 months: The role of vocabulary. *Language Acquisition*, 27(1), 98–115.
- Perkins, L., & Lidz, J. (2021). Eighteen-month-old infants represent nonlocal syntactic dependencies. *Proceedings of the National Academy of Sciences*, 118(41), e2026469118.
- Piaget, J. (1955). *The language and thought of the child*. Meridian Books.
- Pinker, S. (1984). *Language learnability and language development*. Harvard University Press.
- Poesio, M., & Vieira, R. (1997). A corpus-based investigation of definite description use. *arXiv preprint cmp-lg/9710007*.
- Pouscoulous, N., Noveck, I. A., Politzer, G., & Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language acquisition*, 14(4), 347–375.
- Prince, E. F. (1992). The ZPG letter: Subjects, definiteness, and information-status. In W. C. Mann & S. A. Thompson (Eds.), *Discourse descriptions: Diverse linguistic analyses of a fund-raising text* (pp. 295–325). John Benjamins.
- Recanati, F. (2012). *Mental files*. OUP Oxford.
- Rissman, L., Rawlins, K., & Landau, B. (2015). Using instruments to understand argument structure: Evidence for gradient representation. *Cognition*, 142, 266–290.
- Roberge, Y., & Rosen, N. (1999). Preposition stranding and que-deletion in varieties of north american french. *Linguistica atlantica*, 21, 153–168.

- Roberts, C. (2003). Uniqueness in definite noun phrases. *Linguistics and philosophy*, 26, 287–350.
- Roberts, I., & Holmberg, A. (2009). Introduction: Parameters in minimalist theory. In T. Biberauer, A. Holmberg, I. Roberts, & M. Sheehan (Eds.), *Parametric variation: Null subjects in minimalist theory*. Cambridge University Press.
- Rochat, P., Striano, T., & Morgan, R. (2004). Who is doing what to whom? young infants’ developing sense of social causality in animated displays. *Perception*, 33(3), 355–369.
- Ross, J. R. (1967). Constraints on variables in syntax.
- Rozendaal, M. I., & Baker, A. E. (2008). A cross-linguistic investigation of the acquisition of the pragmatics of indefinite and definite reference in two-year-olds. *Journal of Child Language*, 35(4), 773–807.
- Ruff, H. A., & Lawson, K. R. (1990). Development of sustained, focused attention in young children during free play. *Developmental psychology*, 26(1), 85.
- Russell, B. (1905). On denoting. *Mind*, 14(56), 479–493.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *science*, 274(5294), 1926–1928.
- Salomo, D., Graf, E., Lieven, E., & Tomasello, M. (2011). The role of perceptual availability and discourse context in young children’s question answering. *Journal of Child Language*, 38(4), 918–931.
- Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, 69(2), 105–134.
- Saxe, R., Tenenbaum, J., & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10-and 12-month-old infants. *Psychological science*, 16(12), 995–1001.

- Saylor, M. M., & Ganea, P. (2007). Infants interpret ambiguous requests for absent objects. *Developmental Psychology*, 43(3), 696.
- Schaeffer, J., & Matthewson, L. (2005). Grammar and pragmatics in the acquisition of article systems. *Natural Language & Linguistic Theory*, 23(1), 53–101.
- Schafer, R., & de Villiers, J. G. (2000). Imagining articles: What *a* and *the* can tell us about the emergence of DP. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University conference on language development* (pp. 609–620, Vol. 2). Cascadilla Press.
- Schwarz, F. (2013). Two kinds of definites cross-linguistically. *Language and Linguistics Compass*, 7(10), 534–559.
- Seidl, A., Hollich, G., & Jusczyk, P. W. (2003). Early understanding of subject and object wh-questions. *Infancy*, 4(3), 423–436.
- Serratrice, L. (2005). The role of discourse pragmatics in the acquisition of subjects in Italian. *Applied psycholinguistics*, 26(3), 437–462.
- Shi, R. (2014). Functional morphemes and early language acquisition. *Child Development Perspectives*, 8(1), 6–11.
- Shi, R., Cutler, A., Werker, J., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *The Journal of the Acoustical Society of America*, 119(6), EL61–EL67.
- Shi, R., & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science*, 11(3), 407–413.
- Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, 72(2), B11–B21.
- Simons, M. (2013). On the conversational basis of some presuppositions. In *Perspectives on linguistic pragmatics* (pp. 329–348). Springer.

- Skarabela, B. (2007). Signs of early social cognition in children's syntax: The case of joint attention in argument realization in child inuktitut. *Lingua*, 117(11), 1837–1857.
- Skarabela, B., & Allen, S. E. (2002). The role of joint attention in argument realization in child inuktitut. *Proceedings of the 26th annual Boston University conference on language development*, 620–630.
- Sneed, E. (2007). The acquisition of generics. *Evanston, IL: Northwestern University dissertation*.
- Snow, C. E., Pan, B. A., Imbens-Bailey, A., & Herman, J. (1996). Learning how to say what one means: A longitudinal study of children's speech act use. *Social development*, 5(1), 56–84.
- Soderstrom, M., Blossom, M., Foygel, R., & Morgan, J. L. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language*, 35(4), 869–902.
- Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: The relation between action processing and action production in infancy. *Cognition*, 95(1), 1–30.
- Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, 96(1), B1–B11.
- Stalnaker, R. (1974). Pragmatic presuppositions. In N. Munitz & D. Unger (Eds.), *Semantics and philosophy* (pp. 197–219). New York University Press.
- Starkey, P., & Cooper Jr, R. G. (1980). Perception of numbers by human infants. *Science*, 210(4473), 1033–1035.
- Stavans, M., & Baillargeon, R. (2018). Four-month-old infants individuate and track simple tools following functional demonstrations. *Developmental Science*, 21(1), e12500.

- Strawson, P. F. (1950). On referring. *Mind*, 59(235), 320–344.
- Strawson, P. F. (1952). *Introduction to logical theory*. John Wiley & Sons.
- Sugisaki, K., & Snyder, W. (2002). Preposition stranding and the compounding parameter: A developmental perspective. *Proceedings of the 26th annual Boston University Conference on Language Development*, 677–688.
- Sugisaki, K., & Snyder, W. (2006). The parameter of preposition stranding: A view from child english. *Language Acquisition*, 13(4), 349–361.
- Sundara, M. (2018). Why do children pay more attention to grammatical morphemes at the ends of sentences? *Journal of child language*, 45(3), 703–716.
- Suppes, P. (1974). The semantics of children’s language. *American psychologist*, 29(2), 103–114.
- Syrett, K., Kennedy, C., & Lidz, J. (2010). Meaning and context in children’s understanding of gradable adjectives. *Journal of semantics*, 27(1), 1–35.
- Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC neuroscience*, 10, 1–8.
- Theakston, A. L. (2012). “the spotty cow tickled the pig with a curly tail”: How do sentence position, preferred argument structure, and referential complexity affect children’s and adults’ choice of referring expression? *Applied Psycholinguistics*, 33(4), 691–724.
- Therrien, R. (2023). *Anti-locality and preposition stranding in a variety of ontario french* [Doctoral dissertation, Université d’Ottawa/University of Ottawa].
- Tomasello, M., & Haberl, K. (2003). Understanding attention: 12-and 18-month-olds know what is new for other persons. *Developmental psychology*, 39(5), 906.

- Trueswell, J. C., Papafragou, A., & Choi, Y. (2011). Referential and syntactic processes: What develops? In E. A. Gibson & N. J. Pearmutter (Eds.), *The processing and acquisition of reference* (pp. 65–108). MIT Press.
- Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, 73(2), 89–134.
- Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, 40(1-2), 21–81.
- van Hout, A., Harrigan, K., & de Villiers, J. (2010). Asymmetries in the acquisition of definite and indefinite NPs. *Lingua*, 120(8), 1973–1990.
- van Rij, J. (2012). *Pronoun processing: Computational, behavioral, and psychophysiological studies in children and adults* [Doctoral dissertation, University of Groningen].
- Vogels, J., Krahmer, E., & Maes, A. (2013). Cognitive load does not decrease pronoun use when speaker’s and addressee’s perspectives are dissociated. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 35(35).
- Walker, A. (2012). Focus, uniqueness and soft presupposition triggers. *Logic, Language and Meaning: 18th Amsterdam Colloquium, Amsterdam, The Netherlands, December 19-21, 2011, Revised Selected Papers*, 460–469.
- Warden, D. A. (1976). The influence of context on children’s use of identifying expressions and references. *British Journal of Psychology*, 67(1), 101–112.
- Warneken, F. (2016). Insights into the biological foundation of human altruistic sentiments. *Current Opinion in Psychology*, 7, 51–56.
- Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *science*, 311(5765), 1301–1303.

- Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, 59(1), 67–95.
- Wexler, K. (2011). Cues don't explain learning: Maximal trouble in the determiner system. In E. Gibson & N. J. Pearlmuter (Eds.), *The processing and acquisition of reference* (pp. 15–42). MIT Press.
- Williams, A. (2015). *Arguments in syntax and semantics*. Cambridge: Cambridge University Press.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128.
- Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69(1), 1–34.
- Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental science*, 8(1), 88–101.
- Ying, Y., Yang, X., & Shi, R. (2022). Toddlers use functional morphemes for backward syntactic categorization. *First Language*, 42(3), 448–465.
- Yıldız, E., & Uzundağ, B. A. (2022). Mothers' descriptions of referents are related to children's communicative competence. *Cognitive Development*, 64, 101258.
- Zehr, J., & Schwarz, F. (2018). Penncontroller for internet based experiments (IBEX).