

2022-09

Pragmatic Inferences in Preschoolers: Inferring the Speaker's Intended Meaning During Online Comprehension

Ju, Narae

Ju, N. (2022). Pragmatic inferences in preschoolers: inferring the speaker's intended meaning during online comprehension (Doctoral thesis, University of Calgary, Calgary, Canada).

Retrieved from <https://prism.ucalgary.ca>.

<http://hdl.handle.net/1880/115316>

Downloaded from PRISM Repository, University of Calgary

UNIVERSITY OF CALGARY

Pragmatic Inferences in Preschoolers: Inferring the Speaker's Intended Meaning During Online
Comprehension

by

Narae Ju

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PSYCHOLOGY

CALGARY, ALBERTA

SEPTEMBER, 2022

© Narae Ju 2022

Abstract

In this dissertation, I examined children's ability to generate two types of pragmatic inferences during online language comprehension. Using referential communication tasks combined with eye-tracking technology, I assessed children's online pragmatic inferencing in two situations: 1) when inferences are assumed to be context-independent (Chapter 2); and 2) when inferences are closely related to particular aspects of a communicative context (Chapter 3).

In Chapter 2, 4- and 5-year-old children completed a referential communication task with either: (1) a conventional speaker who used language in a conventional way; or (2) an unconventional speaker who was introduced as saying things in a weird way and who used language in an infelicitous and non-optimal way. On critical trials, children were asked to look at an object (e.g., "Look at the big duck") while presented with displays involving a target, a competitor, and either a contrast object (Contrast trials) or another distractor object (No contrast trials). When interacting with a conventional speaker, both 4- and 5-year-old children rapidly generated contrastive inferences upon hearing pronominal size adjectives. That is, children made anticipatory fixations toward the target object when a contrast object existed on the display (e.g., small duck). When there was no contrast object, children showed a tendency to look at the competitor object (i.e., the other big object) during the unfolding adjective. When interacting with the unconventional speaker, children showed a delayed preference to look at the target over the competitor during contrast trials.

In Chapter 3, I examined children's online pragmatic inferencing in a situation that involved knowledge discrepancies between children and a speaker. In Experiment 1, 5-year-old children either possessed privileged knowledge or shared knowledge about object identity with a speaker. Children's tendency to consider a target object (e.g., candle) over a competitor object

(e.g., a candle that looks like an apple) was assessed in a referential communication task. In Experiment 2, children's performance in a condition where the knowledge mismatch was associated with object identity was compared to performance in a condition where the knowledge mismatch related to the physical presence of objects. Results demonstrated children's egocentric bias to consider their privileged knowledge about the actual identities of deceptive objects, rather than taking a speaker's lack of knowledge into account. However, when the knowledge mismatch was created by the awareness of the physical presence of objects, children rapidly considered a speaker's knowledge to guide their interpretation.

Together, these findings demonstrated that preschool children make distinct kinds of inferences to understand a speaker's intended meaning in online comprehension. The efficiency of their online inferencing, however, depends on the specific type of mental representation required in a context. Chapter 4 summarizes and explores the current findings in greater detail.

Preface

While conducting research for this dissertation, the following manuscript was submitted for publication with permission from co-authors S. Graham and C. Chambers. N. Ju did the majority of the writing, was responsible for all data collection and analysis, and was the primary contributor to the manuscript included in this dissertation. All authors provided critical reviews of the manuscripts and contributed to the intellectual content.

Chapter 2 presents as a submitted manuscript to *Child Development* as a revise and resubmit on August 22nd, 2022.

Ju, N., Williams, N., Sedivy, J., Chambers, G.C., & Graham, S.A. (2022). *Preschoolers Modulate Contrastive Inferences during Online Language Comprehension*.

Chapter 3 presents as a submitted manuscript to be submitted to the *Journal of Experimental Child Psychology*.

Ju, N., San Juan, V., Chambers, G.C., & Graham, S.A. (2022). *Five-year-olds' Sensitivity to Knowledge Discrepancies about Object Identity in Online Comprehension*.

Acknowledgements

I am very grateful to my supervisor, Dr. Susan Graham, for her guidance, encouragement, and advice she gave me. I remember the very first day that I joined the Graham Team. Ever since then, Susan has supported me academically and emotionally through the rough journey to finish this thesis. I also would like to express my gratitude to my co-supervisor, Dr. Craig Chambers, for his expertise and insight. He has guided me with invaluable feedback to complete this work. My gratitude also goes to my committee members, Dr. Penny Pexman and Dr. Suzanne Curtin, for their advice and comments throughout the research and the dissertation process. Thank you to Dr. Dimitrios Skordos and Dr. Susan Birch for being a part of my examining committee. I would also like to thank all the committee members for making my defense an enjoyable moment and for your brilliant questions and comments.

This research was supported by the University of Calgary, funding from the Canada Foundation for Innovation, and the Alberta Children's Hospital Foundation awarded to Dr. Susan Graham, as well as an operating grant from SSHRC awarded to Dr. Susan Graham and Dr. Craig Chambers. I am grateful for the financial support that allowed me to pursue my academic career and this research. Many thanks to the children and families who participated in my research. This work would not have been possible without their participation.

Thanks to all the past and present members of Language and Cognitive Development lab, especially Dr. Elizabeth Morin-Lessard and Michelle Zepeda. I really benefited from the brainstorming sessions and conversations with you. Finally, I would like to express my gratitude to my friends and my family. Their constant support and encouragement helped me to navigate the hard times of my PhD life. I could not have made it without you.

Table of Contents

Abstract.....	ii
Preface.....	iv
Acknowledgements.....	v
Table of Contents.....	vi
List of Tables.....	viii
List of Figures.....	ix
CHAPTER ONE: Background and Overview.....	1
1.1 Introduction.....	2
1.2. Online language processing and visual world paradigm.....	3
1.3. Theoretical foundations of pragmatic inferences.....	5
1.4. Context-independent pragmatic inferences: contrastive inferences.....	6
1.5. Context-dependent pragmatic inferences: Perspective-taking.....	9
1.5.1. Theoretical accounts of perspective-taking.....	9
1.5.2. Perspective inferences during online processing.....	11
1.6. Overview of remaining chapters.....	14
CHAPTER TWO: Preschoolers modulate contrastive inferences during online language comprehension.....	17
2.1. Abstract.....	18
2.2. Introduction.....	19
2.3. Method.....	26
2.3.1. Participants.....	26
2.3.2. Apparatus.....	26
2.3.3. Stimuli.....	27
2.3.4. Procedure.....	30
2.4. Results.....	31
2.5. Discussion.....	37
2.6. Supplementary Materials.....	43
2.7. References.....	45
CHAPTER THREE: Five-year-olds' sensitivity to knowledge discrepancies about object identity during language comprehension.....	51
3.1. Abstract.....	52
3.2. Introduction.....	53
3.3. Experiment 1.....	57
3.3.1. Method.....	58
3.3.1.1. Participants.....	58
3.3.1.2. Apparatus.....	59
3.3.1.3. Materials.....	59
3.3.1.4. Procedure.....	63

3.3.2. Data Analyses	66
3.3.3. Results and Discussion	69
3.4. Experiment 2	72
3.4.1. Method	73
3.4.1.1. Participants	73
3.4.1.2. Apparatus and Materials	73
3.4.1.3. Procedure	74
3.4.2. Data Analyses	75
3.4.3. Results and Discussion	76
3.5. General Discussion	78
3.6. References	83
CHAPTER FOUR: Conclusions	91
4.1. Goals and summary of findings	92
4.2. General discussion	94
References	101

List of Tables

Table 1: Chapter 2: List of objects used in critical trials	29
Table 2: Chapter 2: Model summary for the effects of referential context, age, and speaker on target advantage scores	36
Table 3: Chapter 2: Model summary for the effects of referential context and speaker on target advantage scores in each speaker condition	37
Table S1: Chapter 2: Filler descriptions in the unconventional speaker condition	43
Table 1: Chapter 3: List of target objects and three types of competitors	60
Table 2: Chapter 3: Summary of statistical results, Experiment 1	71
Table 3: Chapter 3: Summary of statistical results, Experiment 2	77

List of Figures

Figure 1: Chapter 2: Example of visual display for the "big duck" in contrast trial (a), and no contrast trial (b).....	28
Figure 2: Chapter 2: Relative proportion of fixations to the target versus competitor object the modified noun phrase unfolds in time.....	34
Figure S1: Chapter 2: The proportion of fixations to matching (target and competitor objects) and mismatching (distractors and contrast object) items over time, across critical trials.....	44
Figure 1: Chapter 3: Testing sequence with examples of visual and auditory stimuli	62
Figure 2: Chapter 3: Average target advantage scores within the noun interval, Experiment 1 ..	70
Figure 3: Chapter 3: Example of critical display in Experiment 2 depending on the type of competitor	74
Figure 4: Chapter 3: Average target advantage scores within the noun interval, Experiment 2 ..	76

CHAPTER ONE: Background and Overview

1.1 Introduction

How do we understand what others are talking about? Identifying the objects or events to which a speaker intends to refer is perhaps the most important task for listeners. Otherwise, communication would not succeed. Although this task seems effortless, it is, in fact, not trivial. Consider, for example, the following situation: You are waiting in line for a cupcake and hear a child at the front of the line say, “I will have a small cupcake!”. Upon hearing this statement, you likely start thinking about which size of cupcake you would like to have, as the pronominal adjective *small* implies that there are at least two different sizes of cupcakes available. However, when reaching the front of the line you notice that, in addition to different sizes, there are also a variety of flavours available. Why was the server not confused by the child’s request? You then notice that the child could only see the bottom shelf of the display case, which contained only chocolate cupcakes. As this example illustrates, listeners often *infer* a speaker’s intended meaning using various kinds of non-linguistic information, a process referred to as pragmatic inferencing.

In this dissertation, I examine the pragmatic inferences children generate during online language comprehension. In this chapter, I begin with an overview of the visual world paradigm and online language processing. Next, I outline theoretical frameworks and review two distinct types of pragmatic inferences. I then provide a review of previous research examining children’s online processing for each inference type. In Chapters 2 and 3, I present new research investigating children’s ability to derive pragmatic inferences during the incremental processing of referring expressions. Specifically, in Chapter 2, I describe work examining preschool children’s ability to make *contrastive inferences* in referential interpretation, and whether they can modulate these inferences when warranted. In Chapter 3, I investigate whether children take

a *speaker's perspective* into account in a situation where they either possess privileged knowledge about the identity of objects in the referential context or share that information with the speaker. Finally, in Chapter 4, I provide an overview of my main findings and discuss the results in the context of the existing literature.

1.2. Online language processing and visual world paradigm

Pragmatic inferences involve recovering a speaker's intended meaning, which often goes beyond the semantic meaning of an utterance. To understand how pragmatic inferences are computed during conversations, it is important to consider that conversations take place in a wide range of communicative contexts and a variety of sources of information are spontaneously provided in real-time. Thus, listeners must deal with a significant amount of information to determine a speaker's intended meaning. This issue is in line with questions in the literature of language comprehension around when and how different types of information are integrated. For example, in the domain of syntactic ambiguity, it has been proposed that the initial stage of processing privileges syntactic information, and other types of information such as lexical and pragmatic information are considered at a later stage (Frazier, 1987). More recent views indicate that different types of information can be integrated from the earliest stage of processing. Constraint-based models assume that initial processing does not privilege a particular type of information but that multiple sources of information available in a context immediately influence language comprehension (e.g., Spivey-Knowlton & Sedivy, 1995; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Trueswell, 1996).

To examine how listeners process different types of information in real-time, researchers have used eye-tracking methodologies known as the visual world paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). In the typical setup of this paradigm in a comprehension

study, participants hear an utterance while looking at a display consisting of real or depicted objects. Participants' eye movements are recorded for later analyses, which reflect their visual attention toward the visual input on a display while hearing the linguistic input.

Using this paradigm, researchers have shown, using uninterrupted speech in naturalistic contexts, that language processing is highly incremental (e.g., Tanenhaus et al., 1995). Listeners' eye movements are closely time-locked to the unfolding linguistic signal, meaning that listeners begin fixating denoted objects as a word unfolds. Moreover, listeners make anticipatory fixations to the upcoming referent predicted by the unfolding linguistic input and the visual context (e.g., Altmann & Kamide, 1999; Kukona et al., 2011). These findings suggest that listeners rapidly integrate multiple cues and actively predict the upcoming linguistic item during interpretative processing. Thus, the visual world paradigm allows to examine how listeners use various sources of information during language processing.

Another advantage of the visual world paradigm is that it does not burden participants with extra demands created to make any meta-linguistic judgments (Huettig et al., 2011). Other psycholinguistic tasks, such as lexical decision tasks, require participants to make decisions using a yes/no or word/non-word judgements. In the visual world paradigm, listeners do not necessarily have to make decisions, but rather at the display while hearing spoken sentences. It is thus a powerful tool to investigate children's online language processing. For example, using a simpler version of the visual world paradigm, studies have shown that children around two years of age rapidly associate a familiar target word with a familiar target object after hearing only the partial sound of the target word (Fernald, Swingley, & Pinto, 2001; Swingley, Pinto, & Fernald, 1999). By the age of three, children show incremental processing of adjective-noun phrases (Fernald, Thorpe, & Marchman, 2010; Tribushinina & Mak, 2016).

1.3. Theoretical foundations of pragmatic inferences

Although it seems effortless, understanding a speaker's intended meaning is not so simple. To understand what a speaker intends, a listener must consider why the speaker has decided to use words in the specific way she did in a context where the conversation occurs. The question of how inferences arise has been the subject of intense interest in pragmatics. Grice (1975) advanced a framework of inferences and conversational implicatures, the term often used in psycholinguistics research. A key assumption underlying Grice's theory is that communication is a cooperative effort and rational interlocutors expect each other to comply with this general principle. This principle underlies certain maxims of conversation which state how speakers and listeners should behave regarding the content and the manner of language. Specifically, these maxims are as follows: what is said should be truthful (maxim of quality), as informative as necessary, but not overly so (maxim of quantity), relevant (maxim of relation), clear, unambiguous, brief, and in order (maxim of manner). When listeners apply these maxims to speakers, these expectations allow listeners to generate pragmatic inferences based on the utterance and the context. When such expectations appear to be violated, listeners calculate the extra meaning of the speaker's utterance on the assumption that the speaker is still cooperative and thus purposely doing so. Take, for example, the "I will have a small cupcake!" statement above. The modifier in the utterance would have been over-informative if there was only one size of cupcake in the shop window. By assuming that the speaker still follows the maxim of the quantity or is cooperative at least, listeners would infer that larger sizes of cupcakes must exist, and the speaker does not want one of them.

Grice (1975) proposed a further distinction amongst pragmatic inferences depending on the role of context in the computation of inferences. One type of inference is comparatively context-

independent. That is, some forms of sentences normally convey a certain additional meaning in any context (despite the fact that this meaning is not strictly linguistic in nature), referred to as a *generalized conversational implicature*. In contrast, there are cases where a pragmatic inference is closely related to features of a specific context, and this is referred to as a *particularized conversational implicature*.

1.4. Context-independent pragmatic inferences: contrastive inferences

Context-independent pragmatic inferences have been proposed to arise through the use of specific forms. For example, the utterance *John went into a garden yesterday and found a dog* would typically imply both that John neither owned the garden or the dog. This is because, in the alternative scenario, definite articles such as “the” or “his” would be used. Similarly, a speaker’s use of a prenominal size adjective such as *big* or *small* in descriptions like *the small cupcake* typically implicates that there are other size-contrasting objects from the same semantic category available. This latter example of context-independent pragmatic inference is often referred to as *contrastive inference*.

As the meaning drawn from expressions is pragmatically implied or inferred, it is cancelable either explicitly or contextually. That is, the speaker can provide an additional sentence, which cancels the inferred meaning. The inference can also be canceled when used in a context that does not support the interpretation attached to a certain form of utterance. Thus, although this type of inference occurs across contexts in general, it can be moderated or canceled in a specific situation. The theoretical debate that arises, then, concerns at which point contextual information is taken into account during the interpretation of an utterance. According to Default view, situation-specific features exert effects at a later stage (Levinson, 2000). This view postulates that this type of inference occurs using default mechanisms. For example, contrastive

inferences may be automatically triggered whenever a listener encounters pronominal size adjectives in an utterance. Therefore, pragmatic inferencing cannot be initially blocked. On the other hand, relevance theory (Carston, 1998; Sperber & Wilson, 1986) assumes that pragmatic inferences are constructed by a process through which listeners arrive at an interpretation of an utterance that meets their expectation of relevance with the least cognitive effort. On this account, contextual information is available at all times, which can block the pragmatic interpretation. Other approaches, such as constraint-based approaches (Degen & Tanenhaus, 2019), view contextual information as one of multiple probabilistic constraints that constrain pragmatic inferencing. Depending on how different sources of information are weighted, pragmatic inferences can be moderated.

To summarize, all pragmatic accounts following Gricean framework acknowledge that context-independent inferences can be canceled by contextual information but there is disagreement as to when listeners integrate the contextual information. This raises the question of whether children can cancel or moderate this type of pragmatic inference when the context does not support the pragmatic interpretation. It might be difficult for children to cancel the inferred meaning which has already occurred or to incorporate the contextual information which does not encourage the pragmatic interpretation. Alternatively, it might be easier for children to moderate this type of inference if their expectation that a certain form of an utterance carries certain meaning is not yet as fully developed as that of an adult listener. In Chapter 2 of this dissertation, I assess when and how 4- and 5-year-old children derive contrastive inferences with pronominal size adjectives and how efficiently they adapt their inferencing during online processing, using a visual world paradigm.

To date, only a few studies have examined children's online processing of contrastive inferences. Huang and Snedeker (2013) presented 5-year-olds with critical displays involving either a target object (e.g., big coin) with a contrast object (e.g., small coin) or a target object with another distractor (e.g., button). After hearing prenominal size adjectives (e.g., *Look at the big coin*), 5-year-olds were faster to identify a target object when a contrasting object was present (e.g., a small coin) compared to when there was no contrast object in the display. A recent study further showed that, when given sufficient time to process incoming speech, even 3-year-old children make contrastive inferences when they encounter scalar adjectives in spoken sentences (Davies et al., 2021).

The timing of the contrast effects in children, however, seems different from that in adults. Adult listeners rapidly integrate linguistic and contextual information to make contrastive inferences during the incremental processing of adjectives (Sedivy et al., 1999; Sedivy, 2004; Weber et al., 2006; Grodner & Sedivy, 2011; Wolter, Gorman, & Tanenhaus, 2011). Huang and Snedeker (2013) found that, while adult controls computed contrastive inferences during the unfolding noun, 5-year-olds did so later, after the noun offset. In Davies et al. (2021), young children showed contrast effects when given extra time to process the adjective information by adding a pause after the adjective. Although it has been suggested that children's general cognitive abilities, such as processing speed of language comprehension, might account for the delays in the derivation of contrastive inferences (Davies et al., 2021; Huang & Snedeker, 2013), Davies et al. (2021) failed to find significant correlations between measures of processing speed and contrastive inferencing among children.

To summarize, studies have documented the emergence of children's contrastive inferencing, noting that children's inferences are delayed compared to adults' processing or only

observed when extra time allowed. In these studies, children listened to pre-recorded utterances and were asked to point to an object but did not meet *a speaker*. It is possible that, in the absence of interaction with a speaker, children may not be motivated to actively engage in the task. There is evidence that the interactive nature of a task might facilitate children's developing language skills. For example, in one study, children failed to recruit the semantics of known adjectives to identify an animate referent in a passive looking task, but they succeeded in a more interactive task without time pressure (Syrett et al., 2019). Interactivity is even important for adults in that they will rapidly use speaker-specific information (the identity of a speaker and the experience with the speaker) during the process of interpretation only in interactive settings (Brown-Schmidt, 2009). In the experiment described in Chapter 2, I used a paradigm where children met a speaker and believed that they were receiving spoken instructions from the speaker during the experiment and assessed the time course of children's contrastive inferences.

1.5. Context-dependent pragmatic inferences: Perspective-taking

1.5.1. Theoretical accounts of perspective-taking

As illustrated at the beginning of this chapter, some inferences depend to a greater extent on features of a particular context. When someone says, "The book is good" in a library, interpretation of this utterance is necessarily dependent on the contextual information; otherwise there would be many possible referents available. One way to determine a speaker's intended meaning is to consider a speaker's knowledge, beliefs, or background knowledge, in other words, taking into account a speaker's perspective. In the case of the book example, the listener needs to track what a speaker knows and does not know (e.g., what book they recently read) or which book is mutually visible. Common ground or mutual knowledge consists of the assumptions, knowledge, and beliefs shared among communicative partners (Clark, 2015).

Common ground can be estimated based on different sources of evidence (Clark & Marshall, 1981). These include visual cues (interlocutors are looking at or looked at an object; physical co-presence), linguistic cues (interlocutors talked about an object; linguistic co-presence), and background knowledge (interlocutors are from the same country; community membership).

Although it is broadly agreed that common ground is ultimately considered to resolve reference in assertions and instructions, there is debate in the psycholinguistic literature concerning the timing with which common ground is used during the incremental processing of referential expressions. On one account, the interpretation of referring expressions is restricted to entities which are mutually known to a speaker and a listener (Clark, 1992). When your friend says, *the cupcake*, you can readily exclude the cupcake you ate at home before meeting your friend as a potential referent as your friend would not be aware of it (i.e., this cupcake would be in privileged ground, not common ground). Instead, it would instead be reasonable to assume that referents should be mutually known to you and your friend (i.e., be in common ground). On an alternative account, the initial interpretation of referring expressions will consider all entities regardless of which ground they are located (i.e., common, or privileged ground; Keysar, Barr, Balin, & Paek, 1998). Recruiting the higher level of knowledge about what is in common vs. privileged ground during the moment-by-moment processing might be cognitively effortful. Therefore, this information is not considered from the earliest moments of processing but used later, which in some cases requires the initial interpretation to be corrected.

Researchers have used the visual world paradigm to study the timing with which common ground information is recruited during real-time language processing. In Chapter 3 of this dissertation, I explore whether preschool children immediately consider a speaker's knowledge, whether it is delayed. Furthermore, I investigated whether the efficiency with which

children use a speaker's knowledge in online comprehension differs depending on the type of knowledge mismatch between children and a speaker.

1.5.2. Perspective inferences during online processing

Most experimental studies investigating the role of mutual knowledge in online comprehension have created a discrepancy in visual access to objects on a display between a speaker and a listener. In a typical setup, some objects on a display are hidden from a speaker's view and thus the speaker is not aware of the existence of those objects. In turn, the speaker should then be to refer to the occluded objects in an instruction such as *Pick up the X*. If listeners take the speaker's perspective into account, they will reduce the expectation that the speaker would refer to the objects only available to them (i.e., objects in their privileged ground) and should instead develop the expectation that mutually visible objects (i.e., objects in common ground) would be the possible referents. If, however, listeners determine reference based on the privileged information during the initial stages of language processing, they would consider the object that only they can see (i.e., objects in privileged ground).

Empirical studies of children's use of speaker perspective in online comprehension have shown a somewhat inconsistent pattern of findings. Some (particularly early) studies found that children consider their privileged knowledge of candidate referents during their initial interpretation (Epley et al., 2004; Fan et al., 2015). For example, Epley et al. (2004) presented children aged between 4 and 12 years with a display including, among other things, three objects from the same semantic category but differing in size (e.g., small, medium, and large trucks). When the speaker asked the listener to move *the small truck*, the smallest truck was occluded from the speaker's view and thus the speaker's intended object should presumably be the medium-sized truck from the children's view. During the unfolding description, children tended

to initially fixate to the smallest truck, namely the one that matched the semantics of the instruction but was unknown to the speaker, and in turn, their fixations to the intended object were much delayed.

In contrast, other studies have demonstrated that young children readily take a speaker's perspective into account to guide their interpretation (Khu et al., 2020; Nadig & Sedivy, 2002; Nilsen & Graham, 2009). Nadig and Sedivy (2002) used a display involving a pair of size contrasting objects (e.g., a small and a big glass) along with two distractors. On some trials, one object in the contrast pair was hidden from the speaker's view whereas on the other trials both objects were mutually visible. In this setup, children aged between 5 to 6 years revealed their sensitivity to the speaker's perspective during incremental processing. Although both objects fit the instruction (e.g., "pick up the glass"), children showed a preference to look at the target object over the competitor object when the latter was hidden from the speaker's sight. Thus, when the object in privileged ground had equally good referential fit to another object in common ground, children successfully excluded the referent in privileged ground and considered a speaker's perspective from the early moments of language processing. By contrast, in Epley et al. (2004), the object (smallest truck) in children's privileged ground was the best match for the referring expression (small truck) compared to other trucks in common ground. Since the linguistic fit favoured the object in privileged ground, children were not as successful at inhibiting their privileged ground. When referring expressions were equally matched to potential referents, children flexibly used multiple speakers' perspectives to guide referential interpretation (Khu et al., 2020). That is, 4-year-olds considered two distinct speakers' perspectives to constrain reference from the initial stage of processing, indicating their efficient use of perspective information in a more demanding context.

These mixed findings in the use of a speaker's perspective during incremental processing have also been observed in online comprehension studies with adults, with some studies showing late integration of perspective information (e.g., Barr, 2008; Barr & Keysar, 2002; Keysar et al., 2000; Keysar et al., 2003) and other studies providing evidence against the idea of an "egocentric bias" (e.g., Brown-Schmidt et al., 2008; Hanna & Tanenhaus, 2004; Hanna et al., 2003; Heller et al., 2008; Ryskin et al., 2015). One way to explain the contradictory findings observed in children and adults is that the delayed or immediate effects of perspective-taking are determined by multiple factors, which include the complexity of the referential task, the degree of fit between linguistic information and referential candidates on a display, and listeners' belief about the reliability of mutual knowledge of object existence (Degen & Tanenhaus, 2019; San Juan et al., 2015).

To summarize, the timing of the integration of perspective information seems to depend on the various sources of information available in the context. When the task does not involve too many objects and the referential expression equally fits the objects in common ground and privileged ground, children can take a speaker's knowledge of the presence of objects into account from the earliest moments of processing. One question that arises, however, is whether children can consider other types of perspective inferences in real-time processing. Perspective-taking has been proposed to consist of two levels (Flavell, Everett, Croft, & Flavell, 1981): Level 1 involves an understanding of what others can or cannot see, and the Level 2 requires an understanding of how a mutually visible object can be seen by others. Furthermore, information about a communicative partner's perspective can be acquired through different cues (visual, linguistic, and cultural cues). It is likely that distinct level or type of perspective-taking would

influence when and how children take a speaker's perspective into account during the interpretation of an utterance. This question was addressed in Chapter 3 of this dissertation.

1.6. Overview of remaining chapters

The goal of this dissertation is to provide new insights into children's ability to infer a speaker's intended meaning during incremental referential processing. Across experiments, I use the visual world eye-tracking methodology to acquire fine-grained measures of children's online language processing. These measures help clarify the time course of the pragmatic inferences made during referential processing and the information used in this process.

In Chapter 2, I examined children's context-general inferences during real-time comprehension, with a particular focus on contrastive inferences. Previous studies have shown that preschool children can make contrastive inferences in online comprehension, but that these tend to be delayed compared to adults or are only apparent when speech is slowed down. One goal of my experiment was to replicate and extend these studies. Using visual displays intended to better capture elements of the natural world, I tested whether 4- and 5-year-old children derive contrastive inferences as sentences unfold with natural speech rates. A second goal was to investigate whether preschool children can modulate contrastive inferences when a speaker does not follow the conventional patterns of language use. Although previous literature has shown that adult listeners will reduce their expectations that size adjectives are used to signal contrast when the speaker does not obey the "rules of use" for these expressions (Grodner & Sedivy, 2011; Ryskin et al., 2019), no study so far has investigated the possibility that children flexibly adapt their contrastive inferencing in this way. In this experiment, children heard instructions produced by either a conventional speaker who demonstrated conventional use of language or an unconventional speaker who was introduced as "saying things in a strange way sometimes" and

produced unconventional patterns of language use. If children take their knowledge about a speaker into account, there should be a noticeable difference in the incremental interpretation of size adjectives across the conventional and unconventional speaker conditions.

In Chapter 3, I examined a more context-specific type of inference, namely children's ability to take a speaker's perspective to determine a speaker's intended referent. Across two studies, I test how children manage distinct types of knowledge discrepancies during online comprehension. Research to date has mostly focused on children's communicative perspective-taking in situations that involve a knowledge mismatch regarding object "existence". Typically, these studies use scenarios where child listeners can see objects that are not within the visual field of the speaker (so that the speaker is not aware of them). Although the evidence suggests children can use the speaker's perspective to guide the processing of unfolding referential descriptions, there are questions about the generalizability of these effects. Specifically, it is possible that the ease of using perspective information might differ depending on the type of perspective discrepancy and the associated mental representations. In Experiment 1, I investigated children's tendency to take a speaker's perspective into account in a situation where there is a knowledge mismatch between children and a speaker about the actual identity of objects which have deceptive appearances (e.g., a candle that looks like an apple). In Experiment 2, I examined whether the pattern of children's use of perspective information in real-time comprehension would differ depending on the type of knowledge mismatch. I compared children's tendency to consider their privileged knowledge when this knowledge involves object identity versus object existence. Investigating how children handle different types of perspective discrepancies in this way allows us to have a more comprehensive understanding of children's communicative perspective-taking abilities.

Finally, in Chapter 4, I provide a summary of the findings and relate them to existing theoretical accounts of pragmatic inference. In addition, I describe potential limitations of the studies as well as directions for future research.

**CHAPTER TWO: Preschoolers modulate contrastive inferences during online language
comprehension**

Abstract

This study examined 4- and 5-year-olds' incremental interpretation of size adjectives, focusing on whether contrastive inferences are modulated by speaker behavior. Children ($N = 120$, 59 females, mostly White) encountered either a conventional or unconventional speaker who labeled objects in a correspondingly typical or atypical way. Critical utterances contained size adjectives (e.g., "*Look at the small dog*"). With conventional speakers, gaze measures showed children immediately used the adjective to differentiate members of a contrasting pair, indicating that even 4-year-olds derive contrastive inferences in natural speech. With unconventional speakers, contrastive inferences were delayed in the time-course of processing. The findings demonstrate that preschoolers can adjust their use of pragmatic cues when presented with evidence disconfirming their default assumptions about a speaker.

Keywords: Contrastive inference, Eye tracking, Language comprehension, Adjectival modification, Speaker behavior

2.2. Introduction

Consider the utterance “*Look at the small dog*”, spoken in the presence of a group of dogs. In this case, listeners could begin to identify the speaker’s intended referent upon hearing “*small*”, if the dogs were the only candidates in the context that differed in size. This is because listeners consider the informativity of an unfolding description to make inferences about a speaker’s intended meaning. More specifically, listeners assume a rational speaker would be most likely to use an adjective like “*small*” in the presence of multiple same-category candidates, in order to pick out a specific individual. Previous research has demonstrated that, like adults, child listeners readily generate such *contrastive inferences* upon hearing size adjectives (Huang & Snedeker, 2013; Sedivy et al., 1999). In the present study, we examined 4- and 5-year-olds’ interpretation of size adjectives, with a particular focus on whether children at this age adjust inferences about referential contrast according to whether the speaker follows conventional rules of communication.

To begin, we review the visual world paradigm that has offered insight into the processes underlying pragmatic inferencing in real-time language comprehension (see Huettig et al., 2011 for a review). In typical instantiations of this paradigm, listeners are presented with an array of visual stimuli along with spoken instructions (e.g., “*Look at the duck*”). An examination of listeners’ eye movements during unfolding sentences indicates that both adults and children actively construct a speaker’s intended meaning in real-time by integrating various sources of information during referential processing. For example, adults and children as young as 3 years of age predict a speaker’s intended referent by using linguistic information as well as monitoring the speaker’s visual access to objects (Brown-Schmidt, 2009; Brown-Schmidt et al., 2008; Cane

et al., 2017; Hanna et al., 2003; Heller et al., 2008; Khu et al., 2020; Mozuraitis et al., 2015; Nadig & Sedivy, 2002; Nilsen & Graham, 2009).

In this research, we draw upon the visual world paradigm to examine preschoolers' real-time pragmatic inferences. In particular, we focus on inferences that rely on listeners' default assumptions about patterns of language use, such as the idea that speakers use language optimally and formulate utterances that achieve communicative goals in the most efficient way. Specifically, we focus on the earlier-mentioned case of contrastive inferences derived from prenominal size adjectives (e.g., "*Look at the small dog*"). Here, listeners could rely on a general assumption that a rational speaker will typically design a description using the appropriate amount of information needed to identify an intended referent in the context, avoiding too little or too much information (Grice, 1975). If a speaker must differentiate an object (e.g., a dog) from other size contrasting candidates from the same category, the optimal referential expression will often involve a modified noun phrase (e.g., "*the small dog*"). The size modifier would, however, be redundant in the presence of only a single category member.

Pioneering work by Sedivy et al. (1999) demonstrated that listeners draw on reasoning of this type to guide their interpretation of utterances in real-time. Participants were presented with spoken instructions such as "*Pick up the tall glass*" with an accompanying visual context. On *contrastive* trials, they viewed an array of four objects including two objects that belonged to the same category and differed along a single dimension (e.g., a tall and a short glass). The other objects included an item that could be described using the same adjective (e.g., a tall pitcher) and an unrelated object (e.g., a key). On *non-contrastive* trials, the same-category contrasting object (i.e., short glass) was replaced by an additional unrelated object (e.g., a file folder). When there was a contrast object present, participants derived a contrastive inference and, upon hearing the

adjective (e.g., “*tall*”), anticipated reference to the target object to a greater extent than when there was no contrast object present. Using a similar visual world methodology, Huang and Snedeker (2013) showed that 5-year-olds were faster to locate the target on contrastive trials versus on non-contrastive trials. Further, when given slowed-down spoken instructions, even 3-year-olds show contrastive inferencing during the interpretation of size adjectives (Davies et al., 2021).

In the present study, we pursue the examination of contrastive inferences in 4- and 5-year-olds' processing of naturally-timed speech and investigate whether these inferences have a flexible or automatized character during this developmental period. As background, context-general inferences, such as contrastive inferences, have often been described as conventionalized and default-like in character (e.g., Gazdar, 1979; Horn, 1984). That is, it is proposed that this type of inference is licensed automatically from the speaker's choice of words and the belief that the speaker is obeying conventional rules of language use. For size-based contrastive inferences, the idea is that, when referential contrast is present, listeners will tend to “automatically” assume a size adjective's function is to differentiate the intended object from a same-category alternative. These inferences, however, are not thought to be entirely impervious to contextual factors, in that situation-specific considerations can have an effect, often by moderating or canceling the default inference that arises otherwise (Levinson, 2000).

Questions about the conventionalized nature of contrastive inference have been the focus of adult studies where listeners were presented with speakers who talk in unconventional ways. The reasoning for these studies is as follows: If listeners' expectations about speaker rationality underlie contrastive inferences, listeners should be sensitive to situations where the speaker does not follow expectations and will adjust aspects of real-time comprehension accordingly. For

example, if a speaker uses modifiers in the absence of same-category alternatives, listeners may reduce or suspend contrastive inferencing because the modifiers are in turn redundant informative. If, however, contrastive inferences are more automatized or context-invariant, the behavior of a situation-specific speaker may have no or little influence on listeners' interpretation. Using these paradigms, studies have shown that adult listeners can adjust their contrastive inferences when given evidence that highlights the nonconventional communicative behavior of the speaker (Gardner et al., 2021; Grodner & Sedivy, 2011; Ryskin et al., 2019). However, to produce this effect, it appears that the relevant evidence needs to be both strong and abundant, including the speaker's use of over- and under-informative descriptions, incorrect labels, and often explicit statements about the speaker's language impairment (Grodner & Sedivy, 2011; Ryskin et al., 2019).

One explanation for the robustness of adults' contrastive inferences involves highly ingrained expectations about language use built from a lifetime of experience with rational speakers (Ryskin et al., 2019). It is therefore possible that children may be more likely to adapt their contrastive inferences based on speaker behaviour. This is because their more limited experience should entail a less entrenched expectation that a size adjective will signal referential contrast. Further, young children may attend to different aspects of behavior than adults do or may weigh this information differently in communicative interactions. For example, whereas adults can readily draw on their understanding of the core lexical meaning of words, children's lexical knowledge is more limited. Therefore, they might be more sensitive to a broad variety of speaker behaviors in assessing a speaker's referential intent. These include the reliability of a given speaker (such that past "mislabeling" makes children reluctant to learn new word-referent associations from that individual: e.g., Henderson et al., 2015; Koenig & Harris, 2005; Scofield

& Behrend, 2008) as well as a variety of emotional cues produced by speakers (e.g., Berman et al., 2013; Graham et al., 2006). This broad sensitivity could increase their ability to adjust contrastive inferencing in response to an individual who fails to follow conventional rules of conversation.

Alternatively, a second possibility is that, like adults, children might show a strong contrast effect even when a speaker does not follow the conventional use of a language. Young children, even 3-year-olds, can derive contrastive inferences from prenominal size adjectives during online comprehension (Huang & Snedeker, 2013; Davies et al., 2021). Furthermore, evidence suggests that the frequency of adjectives in child speech increases from two to three years of age and then plateaus (Tribushinina & Gillis, 2012). In child-directed speech, contrastive adjectives are more frequently used to describe a referent relative to a contrasting object (Davies et al., 2020). As such, it is possible that the process of deriving contrastive inferences quickly becomes automatized during early childhood.

Our specific goals were as follows: First, we sought to replicate previous findings demonstrating children's ability to make contrastive inferences during online comprehension with 4- and 5-year-olds. Although recent evidence suggests that even 3-year-olds can derive contrastive inferences, the effect is not observed with natural speech rates. Instead, the effect is found only when the spoken instruction is slowed down, and a pause is inserted between the adjective and the target noun (Davies et al., 2021). The authors suggested that the inability of 3-year-olds to engage in in-the-moment contrastive inferencing likely arises from young children's processing limitations. That is, young children seem to have the pragmatic skill to infer a speaker's meaning in online comprehension, but its manifestation is hindered when they do not have sufficient time to process all the necessary information. Given the rapid development of

children's cognitive abilities during the preschool years, we predict that 4-year-olds will be more successful at deriving contrastive inferences when modified noun phrases are encountered in fully fluent speech spoken at a natural rate.

The second goal was to investigate the extent to which contrastive inferences can be influenced by child listeners' in-the-moment beliefs about the communicative rationality of the speaker, or instead if these inferences are largely insensitive to this kind of situation-specific factor. Work on parallel topics, such as the effect of momentary disfluencies on referential processing, shows that children can flexibly adjust online processing in response to knowledge about the speaker's traits and/or speaker behavior (Orena & White, 2015; Thacker et al., 2018a, 2018b). Drawing on a paradigm first used by Grodner and Sedivy (2011), the design of the present study was such that child listeners interacted with either a conventional or an unconventional speaker. We implemented speaker conventionality as a between-subjects factor to provide children with straightforward evidence that the speaker either obeyed or broke conversational rules. Children in each speaker condition encountered critical trials in which referential contrast was either present or absent. On *contrast* trials, the visual display consisted of a target object that matched the modified noun phrase (e.g., big duck, given the description "*the big duck*"), a competitor object that matched the target object in size but came from a different category (e.g., big hippo), an unrelated object (e.g., bee), and a contrast object (e.g., small duck). On *no contrast* trials, the contrast object was replaced with an unrelated object (e.g., coin). We measured which objects children considered upon hearing the modified noun in the instruction ("*Look at the big duck*").

The manipulation varying speaker conventionality was somewhat different than past studies. In previous research, the unconventional speaker produced various types of nonoptimal

sentences: over-informative descriptions (i.e., where adjectives were unnecessary), under-informative descriptions (i.e., where the lack of modifiers led to referential ambiguity), and sometimes incorrect descriptions (i.e., where the labeling of object or location was wrong; Gardner et al., 2021; Grodner & Sedivy, 2011; Ryskin et al., 2019). Under-informative and incorrect descriptions provided by a speaker might give children the impression that the speaker is unreliable and inaccurate instead of unconventional in language use. The present study incorporated a different type of unconventional language use. More specifically, the unconventional speaker always provided sufficient information to allow the child to identify the intended referent, but this information was sometimes delivered in atypical and nonoptimal ways. For example, in the unconventional speaker condition, a banana might be described as “*It’s a long yellow fruit you can peel, that monkeys like to eat*”. The conventional speaker, on the other hand, would simply use a conventional noun (“*It’s a banana*”). This helped ensure children did not become frustrated with the task or the speaker, and provided evidence about the speaker’s communicative oddness. Other work has, in fact, shown that evidence of this sort can affect children’s real-time interpretation of referring expressions. For example, a speaker who expresses unconventional opinions (e.g., says a rotten banana “looks yummy”) leads children to adjust the way they use that speaker’s emotional prosody to determine referential intent (Thacker et al., submitted). Given these findings, we predict that when presented with an unconventional speaker children will be less likely to draw contrastive inferences as a prenominal adjective is encountered.

2.3. Method

2.3.1. Participants

Children were recruited from a participant database in a large city in Western Canada. Data from 64 4-year-olds (34 boys, $M_{age} = 4.68$ years, $SD = .12$, range = 4 years 6 months to 4 years 11 months) and 56 5-year-olds (27 boys, $M_{age} = 5.77$ years, $SD = .17$, range = 5 years 6 months to 5 years 11 months) were included in the final sample. Each child was assigned to one of the two speaker conditions: the conventional speaker condition (31 4-year-olds and 30 5-year-olds), and the unconventional speaker condition (33 4-year-olds and 26 5-year-olds). This sample size is sufficient to detect a medium effect size with 95% power in a conventional repeated-measures analysis of variance (ANOVA) for a relevant interaction with alpha at .05 (Faul et al., 2007). An additional 34 children were tested but were excluded from the analyses for the following reasons: failure to meet the language criteria (less than 80% of exposure to English; $n = 10$), failure to complete the task ($n = 3$), calibration failure ($n = 1$), prior participation in the same or similar task ($n = 3$), insufficient eye gaze data ($n = 16$), or experimenter error ($n = 1$). All children in the final sample used English as their primary language. Most parents identified their child as European-Canadian (86%) and had post-secondary or higher education (90%). This study was approved by the institution's research ethics board, and informed consent was obtained from the children's parents.

2.3.2. Apparatus

The visual stimuli were presented on a 46-inch screen. Children's eye movements were tracked using a Tobii Pro x3-120 table-mounted eye tracker located on a small table in front of the children and underneath the screen. The experiment was conducted using E-prime extensions for Tobii and was preceded by a standard 5-point calibration routine. Areas of interest (AOIs) were

defined prior to data collection based on the four shelf compartments (i.e., top right, top left, bottom right, or bottom left). Children's gaze position was logged every 8.3 ms, and fixations were defined as stable looks to a location that lasted longer than 95 ms.

2.3.3. Stimuli

The test trials consisted of 32 different visual displays (16 critical trials and 16 filler trials). In each display, there was a white shelf unit with four compartments. Each compartment contained a distinct familiar object. All images were created from pictures of either actual objects (e.g., coin) or toy versions of the objects (e.g., rubber duck for duck). Shadow effects were applied to each image to make the object look more realistic and highlight the relative differences in size across objects.

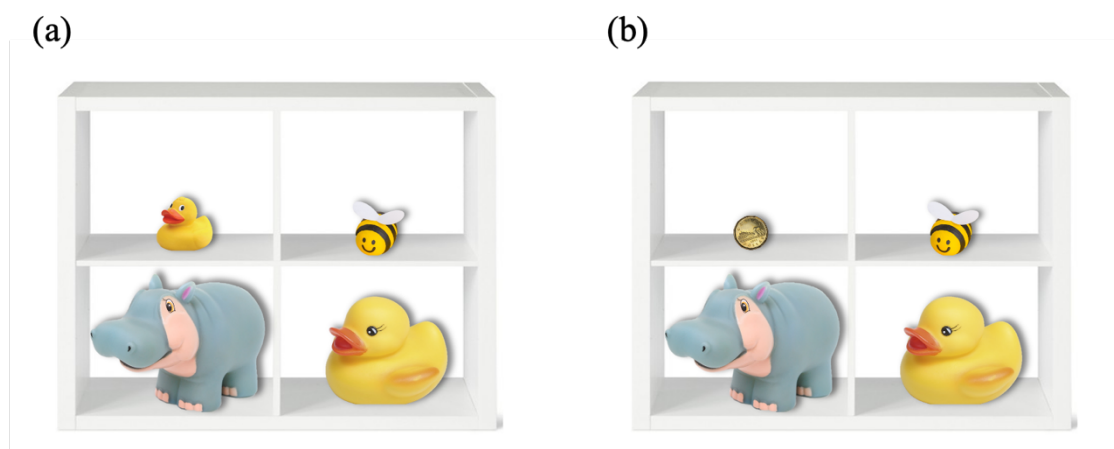
The auditory stimuli presented with the visual display on each trial were pre-recorded by a female native English speaker. The speaker also recorded the utterances used in **trait-revealing** trials that established the (un)conventionality of the confederate speaker. On these trials, a banana and a telephone were presented one at a time on the screen. In the conventional speaker condition, these items were referred to using an unadorned noun (e.g., "*It's a banana.*"), whereas in the unconventional speaker condition an over-informative and elaborate description was used to refer to the same object (e.g., "*It's a long, yellow fruit you can peel that monkeys like to eat.*").

Test trials followed the trait-revealing trials. See Figure 1 for an example of the critical trial displays as a function of the referential context and Table 1 for a list of all objects presented on critical trials. On critical trials, the visual display consisted of four objects: a target object (e.g., big duck), a competitor object that shared the dimensional property of the target but belonged to a different category (e.g., big hippo), an unrelated distractor object (e.g., bee), and, critically, an object whose identity was varied to create the referential context manipulation. This

object was either an object from the same category as the target but differing in size (e.g., small duck) or another unrelated distractor (e.g., coin). Further, the competitor object was adjusted to be slightly smaller or bigger than the target object such that it would provide a better match for the size adjective on purely semantic grounds. For a given participant, half the critical trials included a same-category contrast for the target referent (contrast condition; e.g., a big duck and a small duck), whereas the other half had no contrasting alternative present (no contrast condition; all four objects belonged to distinct categories). The pairing of displays to referential context conditions (contrast vs. no contrast) was counterbalanced to produce two test orders. The referential instructions in critical trials were always the form of the phrase “*Look at the [size adjective: big or small] [target object]. Point to it.*” in both speaker conditions.

Figure 1

Example of visual display for the “big duck” in contrast trial (a), and no contrast trial (b)



Note. The images are recreated for illustrative purposes only. The visual displays used in the studies are similar but not identical to this figure.

Table 1*List of objects used in critical trials*

Item	Adjective	Target	Competitor	Distractor 1	Distractor 2*
1	Big	Duck	Hippo	Bee	Coin
2	Big	Glass	Present	Cherry	Bird
3	Big	Crayon	Monkey	Fish	Peanut
4	Big	Hat	Potato	Block	Lamp
5	Big	Elephant	House	Strawberry	Camera
6	Big	Cookie	Clock	Butterfly	Fork
7	Big	Donut	Penguin	Lemon	Apple
8	Big	Shoe	Laptop	Button	Bowl
9	Small	Helmet	Rabbit	Dress	Box
10	Small	Dog	Candy	Jar	Blanket
11	Small	Cup	Key	Book	Horse
12	Small	Tomato	Ring	Whale	Bucket
13	Small	Chocolate	Blueberry	Puzzle	Dinosaur
14	Small	Doll	Egg	Bag	Giraffe
15	Small	Pig	Cracker	Truck	Umbrella
16	Small	Car	Turtle	Robot	Pillow

Note. Asterisk indicates an unrelated object used in no contrast trials. On contrast trials, there was a contrast object (e.g., small duck) instead of distractor 2.

On filler trials, the visual displays consisted of two comparatively small and two comparatively big objects, with two of the objects belonging to the same superordinate level category (e.g., a small notebook and a small pencil) and two objects from different categories (e.g., a big hamburger and a big sheep). The instructions were used to remind participants of the speaker's status as a conventional language user, such that the conventional speaker used unmodified noun phrases to refer to the target object (e.g., "*Look at the notebook*"), whereas the

unconventional speaker provided elaborate descriptions of the target object, similar to the example given in the trait-revealing trials (e.g., “*Look at thing made of paper that you can write or draw in*”). Descriptions provided by the unconventional speaker in filler trials are presented in Supporting Information. The target object was one of the two members of the same superordinate category in half of the filler trials, and one of the objects outside of this category in the other half. The size of the target object was also counterbalanced across filler trials.

2.3.4. Procedure

The experimenter first introduced the child to a confederate in a small room adjacent to the testing room. The confederate, “Monica”, was seated in front of a computer screen wearing a headset microphone. The child was told that Monica would play the game from this room, and she would be able to talk to and hear the child in another room through her headset. The confederate did not say anything but smiled and waved her hand. The purpose of the confederate was to suggest to children that they were interacting with a live speaker to make the task seem more naturalistic. In fact, all auditory stimuli were pre-recorded to ensure uniformity in speech rate and intonation.

After meeting the confederate, the child was led to the testing room and was seated in a small chair in front of the eye tracker. Using Tobii Clearview Software, children’s eye gaze was first calibrated. This step required at least three of five fixation points to be successfully registered. The experimenter then provided different descriptions about the speaker depending on the speaker condition to which children were randomly assigned. Children in the conventional speaker condition were told several neutral facts about the confederate (i.e., “Monica has dark hair and dark eyes. She likes to dance and play sports. She also likes to read books and go to the movies.”) whereas those in the unconventional speaker condition were told that the speaker

communicates in an unusual way (i.e., “Monica sometimes says things in a very strange way. She says things in a different way than you or I would. Yes, she says things in a strange way.”). This description was then followed by two trait-revealing trials. Children were presented with a familiar item on the screen and asked to name it (e.g., a banana). The experimenter (in the room with the child participant), who was also wearing a microphone headset, then pretended to ask the confederate what she would call the object. In the conventional speaker condition, children heard conventional referential expressions (e.g., “*It’s a banana.*”). Children in the unconventional speaker condition heard overly informative and infelicitous referential expressions (e.g., “*It’s a long, yellow fruit you can peel that monkeys like to eat.*”).

After the trait-revealing trials, participants completed 32 test trials. They were asked to listen to and follow the speaker’s instructions. The experimenter in the room with the child pretended to interact with the confederate speaker, calling the confederate’s name and letting her know that participants were ready at the beginning of each trial. On each trial, the speaker’s instruction always began about one second after the visual display appeared. For each of the two test orders, critical and filler trials were presented in a pseudorandomized order, such that the same type of trial was not presented more than two consecutive times. The location of the target object was counterbalanced across test trials.

2.4. Results

Eye movement measures were used to capture children’s real-time understanding of the speaker’s referential intent as the noun phrase was heard, and how this understanding was modulated by context (contrast vs. no contrast) and by speaker conventionality (conventional vs. unconventional). The R code and supplementary analyses are available at https://osf.io/uhnky/?view_only=f3bfbe453c974810baf9a2aa1a330531. We began by first

screening for poor eye-gaze data on individual trials using the eyetrackingR package (Dink & Ferguson, 2018). We defined poor eye gaze-data as high level of track loss or non-AOI gaze. Non-AOI gaze occurred when children looked at the display on the screen, but their point-of-gaze did not fall within one of the AOIs. Critical trials in which track loss occurred in more than 70% of the entire recorded interval of the critical instruction were removed from the analysis. Note that developmental eye tracking studies often use a 50% track loss criterion. We used a 70% criterion because we included non-AOI gaze as track loss. The main findings were consistent regardless of the track loss criterion used. Results based on a 50% track loss criterion are provided in the supplemental material. Participants for whom more than three trials had been excluded from one of the contrast conditions (contrast vs. no contrast) were entirely excluded from the analysis to ensure data integrity. For the remaining participants, the track loss criterion led to 5.4% of trials being removed from the analysis.

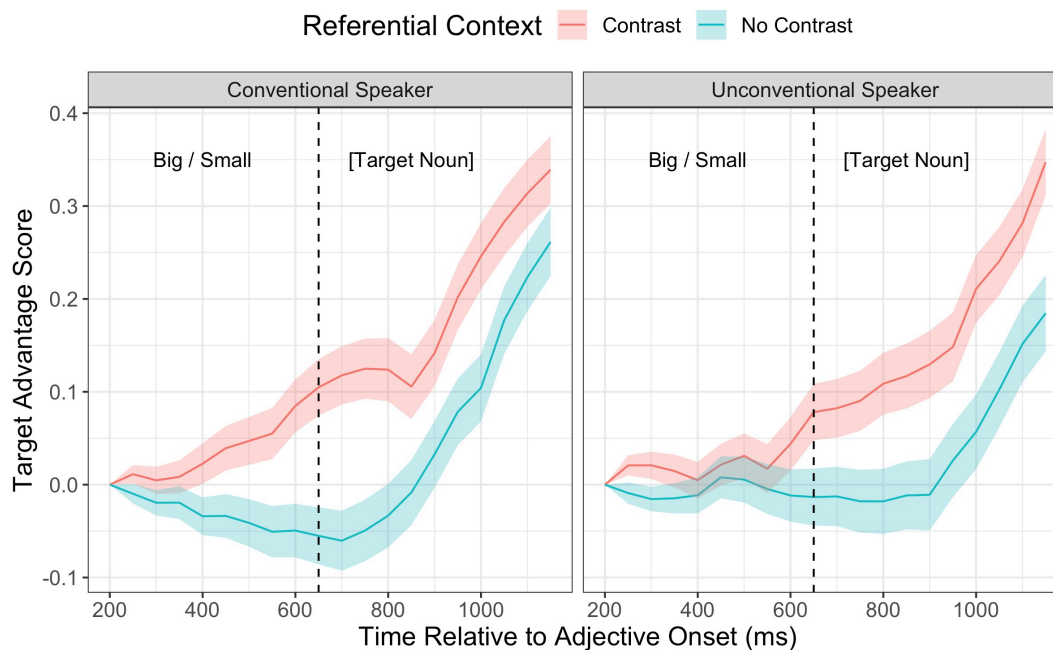
Our key analyses focused on fixations upon hearing the adjective in the instruction (e.g., “big” in “big duck”). The duration of the critical time window was 450 ms, which reflected the average duration of the adjective across trials (440 ms). A 200 ms margin was added to the onset and offset of this time interval, reflecting the time needed for eye movements to be programmed and executed in response to perceptual input (Matin et al., 1993; Trueswell, 2008). We removed any fixations initiated before the left margin to ensure the data reflect gaze patterns driven by the content of the target expression and were not lingering fixations launched before the adjective was heard.

We first aggregated the data into 50 ms time bins and calculated the proportion of fixations to the target and competitor objects by dividing their respective fixation measures by the sum of all fixations to all objects. The proportion of fixations to items matching the adjective

(target and competitor objects) and items mismatching the adjective (distractors and contrast object) are presented in Supporting Information. For the analyses, target advantage scores were calculated using the proportion of fixations to the target and the competitor objects (average proportion of fixations to target minus average proportion of fixation to competitor). This measure produces a single value reflecting the extent to which participants considered the target object relative to the competitor object. Figure 2 shows average target advantage scores across the speaker and referential context conditions. Given the lack of age effects in the subsequent statistical analyses, fixations are collapsed across age groups. Note that positive scores indicate greater visual consideration of the target object, and negative scores indicate greater consideration of the competitor object. Visual inspection of the graph indicates that, upon hearing the adjective, children in the conventional speaker condition were more likely to look at the target object in contrast trials and the competitor object in no contrast trials. In the unconventional speaker condition, however, children's tendency to fixate to the target object compared to the competitor object is delayed on contrast trials. Further, children in this condition did not show a preference to fixate to the target over the competitor object or vice versa while processing the adjective on no contrast trials.

Figure 2

Relative proportion of fixations to the target versus competitor object the modified noun phrase unfolds in time.



Note. Scores above 0 indicate more fixations to target object, and below 0 indicate more fixations to competitor object.

Fixation data were analyzed in R (R Core Team, 2020) using the lme4 (Bates et al., 2015) and lmerTest packages (Kuznetsova et al., 2017). Target advantage scores were rescaled between 0 and 1 and transformed using a logit transformation for the analyses. We used growth curve analysis (GCA; Mirman, 2014) to explore the effect of speaker conventionality and referential context on the time course of referential processing within the critical window. Fixation patterns over time were modeled with second-degree orthogonal polynomials (including linear and quadratic) as well as fixed effects of age (4-year-olds vs. 5-year-olds), speaker (conventional vs. unconventional), and referential context (no contrast vs. contrast). The linear term reflects steady growth in the target advantage score, and the quadratic term reflects patterns where the rate of

change shifts. All fixed effects were contrast coded (.5 vs. -.5). For the random effects structure, the model allowing convergence without a singular fit warning included random intercepts of participants and items. To determine the best fitting model, we compared the linear model against a quadratic model using the ANOVA function in R. The linear model provided a better fit for the data, $\chi^2(8) = 8.32, p = .402$.

The model estimates for the linear model are shown in Table 2. There was a significant effect of referential context, reflecting the overall greater consideration of the target object when there was a same-category contrast object in the display compared to when this object was replaced with an object from a different category ($\beta = -.40, SE = .05, p < .001$). The effect of referential context was also significant on the linear time term ($\beta = -1.11, SE = .22, p < .001$), reflecting the overall greater tendency to fixate to the target object on contrast trials. No effect involving age was significant, except a significant interaction between referential context, speaker, and age ($\beta = .39, SE = .20, p = .048$). Further analyses conducted using the emmeans package (Length, Singmann, Love, Buerkner, & Herve, 2019) revealed a contrast effect across the speaker conditions for both 4- (conventional speaker: $\beta = -.16, SE = .05, p = .001$; unconventional speaker: $\beta = -.10, SE = .05, p = .038$) and 5-year-olds (conventional speaker: $\beta = -.36, SE = .05, p < .001$; unconventional speaker: $\beta = -.14, SE = .05, p = .009$). Thus, the significant interaction showed a significant contrast effect in each speaker condition at both age groups, but the contrast effect was the largest in the conventional speaker condition among the older children.

Table 2

Model summary for the effects of referential context, age, and speaker on target advantage scores

Effect	<i>Estimate</i>	<i>SE</i>	<i>t</i>	<i>p</i>
(Intercept)	0.07	0.11	0.61	.549
Linear	0.23	0.32	0.72	.480
Context	-0.40	0.05	-8.20	<.001
Age	0.08	0.07	1.26	.210
Speaker	-0.02	0.07	-0.33	.741
Linear:Context	-1.11	0.22	-5.05	<.001
Linear:Age	0.31	0.22	1.42	.157
Context:Age	0.16	0.10	1.69	.091
Linear:Speaker	-0.01	0.22	-0.03	.978
Context:Speaker	-0.37	0.10	-3.80	<.001
Age:Speaker	-0.12	0.13	-0.89	.373
Linear:Context:Age	0.22	0.44	0.51	.609
Linear:Context:Speaker	-1.19	0.44	-2.73	.006
Linear:Age:Speaker	-0.44	0.44	-1.00	.316
Context:Age:Speaker	0.39	0.20	1.98	.048
Linear:Context:Age:Speaker	1.13	0.87	1.29	.198

Most relevant to our hypotheses, we found significant interactions between referential context and speaker ($\beta = -.37$, $SE = .10$, $p < .001$) and between referential context, speaker, and the linear time term ($\beta = -1.19$, $SE = .44$, $p = .006$). To explore the interaction patterns, we conducted separate analyses for each speaker condition (see Table 3). Children in both speaker conditions directed reliably more fixations overall to the target object on contrast trials compared to no contrast trials (conventional speaker: $\beta = -.57$, $SE = .07$, $p < .001$; unconventional speaker:

$\beta = -.22, SE = .07, p = .001$). However, the interaction of referential context and the linear time term was significant only in the conventional speaker condition ($\beta = -1.67, SE = .31, p < .001$). In this condition, the presence of a contrast object led children to consider the target object as the adjective was heard. When the contrast object was absent, children instead tended to fixate to the competitor object upon encountering the adjective. In contrast, when interacting with an unconventional speaker, children's tendency to fixate to the target object over the competitor was delayed on contrast trials, leading to the absence of a reliable contrast effect within this condition alone.

Table 3

Model summary for the effects of referential context and speaker on target advantage scores in each speaker condition

Effect	Estimate	SE	t	p
Conventional Speaker				
(Intercept)	0.06	0.13	0.46	.648
Linear	0.24	0.37	0.66	.518
Context	-0.57	0.07	-8.23	<.001
Linear:Context	-1.67	0.31	-5.37	<.001
Unconventional Speaker				
(Intercept)	0.08	0.07	1.11	.272
Linear	0.21	0.15	1.40	.162
Context	-0.22	0.07	-3.19	.001
Linear:Context	-0.50	0.30	-1.65	.099

2.5. Discussion

The present study investigated 1) whether 4- and 5-year-old children make contrastive inferences in real-time when interpreting size adjectives and 2) whether speakers who differed in

their adherence to communicative conventions shifted these inferences. Our findings offer key insights into the interaction between speaker behaviour and pragmatic inferencing during early childhood.

First, upon hearing a prenominal size adjective, both 4- and 5-year-olds anticipated that the speaker was referring to one member of a contrasting pair, which is largely consistent with previous research (Davies et al., 2021; Huang & Snedeker, 2013). Notably, the current findings showed that, unlike 3-year-olds who show online contrastive inferencing only when given extra processing time (Davies et al., 2021), 4-year-olds can rapidly compute contrastive inferences in naturally-paced fluent speech, with effects emerging during the unfolding adjective. This suggests that, by 4 years of age, children can evaluate linguistic input against the referential context and draw relevant pragmatic inferences that affect their interpretation of utterance information in real-time.

What might account for the difference between our (Experiment 1) results and those of Davies et al. (2021), aside from the contribution of age-related developmental changes? In our study, we used visual stimuli that looked more realistic (in terms of shading and depth cues), which may have helped children distinguish relative size differences between objects. Furthermore, our task was interactive in that children met a confederate speaker and believed that they were communicating with her from the other room. As noted by Davies et al. (2021), young children's emerging pragmatic competence can be revealed more effectively when factors that make tasks more challenging have been removed (Falkum et al., 2017; Pouscoulous & Tomasello, 2020; Schulze et al., 2013; Stiller et al., 2015; Syrett et al., 2019). The methodological features of the present study might have facilitated pragmatic reasoning among young children. Nevertheless, the present results add to the existing literature documenting

developmental continuity of contrastive inferences among preschool children by showing that children around 4 years of age derive contrastive inferences in spontaneous speech.

Second, our findings demonstrated that children modulated their expectations about the role of pronominal adjectives based on the speaker's communicative behaviour. When interacting with a speaker who violated conventions of language use, the immediacy of children's contrastive inferences was diminished, leading the preference to fixate to the target object over the competitor object to emerge later in course of processing. These results suggest that, like adults, young children can modulate pragmatic inferences depending on their knowledge of the speaker (Gardner et al., 2021; Grodner & Sedivy, 2011; Ryskin et al., 2019). However, in contrast to the strong manipulations used in adult studies (in which an unreliable speaker produced both over- and under-informative sentences), the unconventional speaker in the present study did not use modifiers in an incomplete manner but departed from communicative norms by using long-winded and infelicitous descriptions for objects that could be described with an unadorned noun. Earlier, we noted two factors that could explain children's sensitivity to speaker conventionality in the context of contrastive inference, namely weaker experience-based expectations for size adjectives to signal referential contrast, and a greater sensitivity to speaker behaviors when inferring communicative intent. Although both of these may play a role, it is interesting to note that (consistent with the second explanation) the findings of the present study share parallels with the way in which preschool children adjust their use of disfluency cues or emotional prosody in language processing in response to a variety of speaker behaviors (e.g., Orena & White, 2015; Thacker et al., submitted). It is also worth noting that the present results showed that the effect of speaker unconventionality was specific to pragmatic inferencing and was not an overall suspension of incremental interpretation. Children in both conditions showed

a trend to look more to the competitor object than the target object when there was no size contrasting alternative on the display (i.e., in the no contrast trials), indicating that they continued to process the semantics of the size adjective (recall that the competitor object was a better fit for the size adjective in absolute terms).

Together, the present findings suggest that 4- and 5-year-old children integrate information about a speaker with their prior expectations about the contrastive function of prenominal size adjectives to make pragmatic inferences. Despite the effect of speaker conventionality on contrastive inferences, we found that, in absolute terms, children in the unconventional speaker condition still made more fixations to the target object relative to the competitor object when there was a size contrasting object during the unfolding adjective. One possible explanation is that this finding relates to the automatized processing of contrastive inferences. As noted earlier, previous studies have shown that adult listeners still derived contrastive inferences upon hearing descriptions with size adjectives, even after exposure to a vast number of examples of a speaker using adjectives infelicitously and other evidence of the speaker's pragmatic and linguistic incompetence (Ryskin et al., 2019). The fact that contrastive inferences were delayed, but not fully suppressed, may suggest that, even by 4 years of age, children have accumulated substantial experience with a rational speaker using pragmatically-motivated prenominal adjectives (i.e., used in the presence of referential contrast), leading to comparatively "entrenched" patterns of interpretation. It is also worth noting that complete suppression of contrastive reasoning might be unlikely based on the unconventional speaker's behavior. Although the speaker departed from communicative norms by producing unnecessarily long and convoluted descriptions, this alone does not indicate she is incapable of producing pragmatically felicitous contrastive descriptions.

The fact that children show adult-like automatized processing of contrastive inferences connects with a recent word learning model proposed by Bohn et al. (2021). This account proposes that adults and children possess similar processing architectures in which multiple sources of information are integrated to infer the meaning of novel words, but what changes over time is the degree of sensitivity to each information source. Within this framework, it is likely that adults and children also assign a different weight to distinct information sources during pragmatic inferencing. That is, adults might have increased sensitivity to the typical contrastive function of size adjectives while children's sensitivity to this information source is still developing. In this case, speaker unconventionality would exert a stronger influence on incremental processing among children compared to adults. It remains an open question to what extent the information-integration model in language learning can be straightforwardly applied to the online integration of information sources during language.

The present findings suggest a number of key directions for future research. First, it would be fruitful to assess whether adult listeners can adapt their pragmatic expectations when presented with speaker unconventionality that does not include incomplete or incorrect utterances. As noted earlier, previous research has suggested that adults adjust their contrastive inferences when provided with strong and abundant evidence that the speaker is unconventional (i.e., use of over- and under-informative descriptions, incorrect labels, and explicit statements about the speaker's language impairment: (Grodner & Sedivy, 2011; Ryskin et al., 2019)). It remains to be determined whether the unconventional labeling used in this study (i.e., verbose description) would similarly disrupt adult listeners' online contrastive inferencing. A second direction for future research is to examine whether the current results could be generalized to other populations. The data for the present study was collected from a western, educated,

industrialized, rich, and democratic (WEIRD, Henrich et al., 2010) society. Thus, future research should examine pragmatic reasoning and adaptation in children from diverse populations. We also hope more studies will be conducted with clinical populations, such as children who have difficulties in social communication (e.g., individuals with social communication disorder). Experimental work along these lines is essential for understanding the universality of the mechanisms underlying pragmatic processing.

In summary, the findings of the present study provide evidence that, during real-time language comprehension, 4- and 5-year-old children spontaneously apply pragmatic inferences reflecting default assumptions about rational communication. Further, we found that child listeners adjust these inferences in response to the speaker's adherence to communicative norms.

2.6. Supplementary Materials

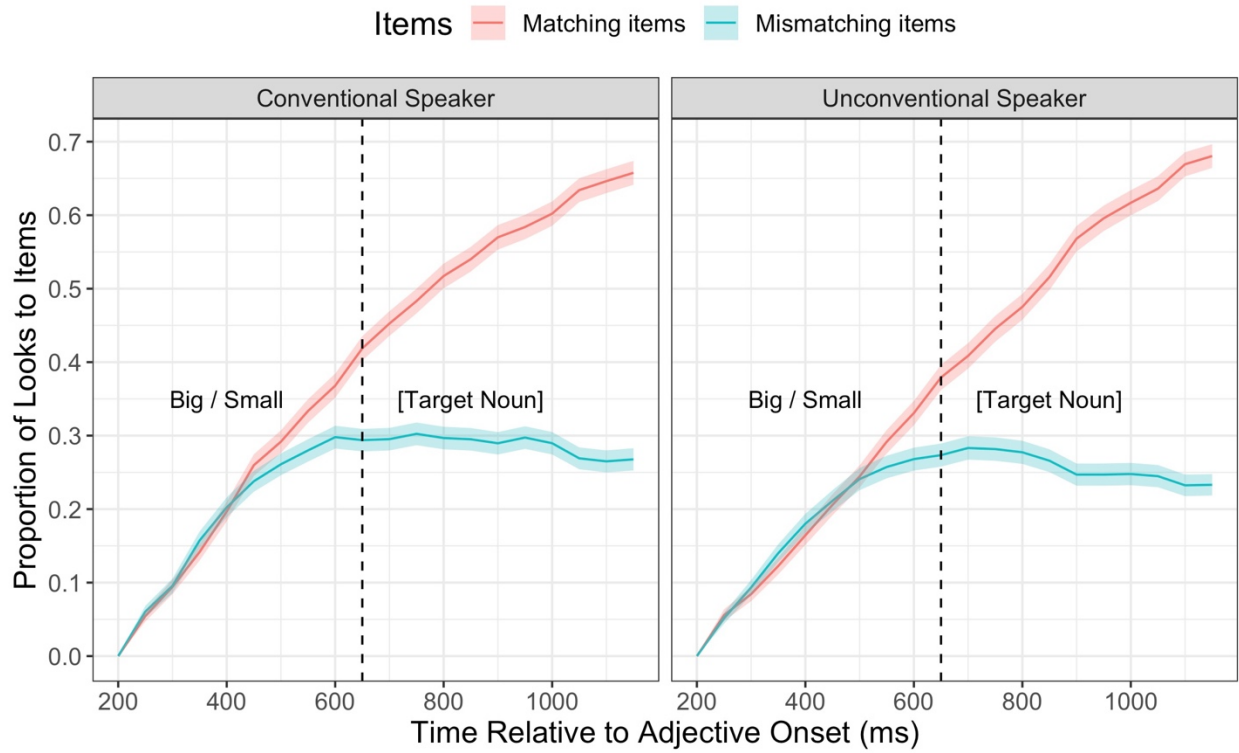
Table S1

Filler descriptions in the unconventional speaker condition

Target object	Descriptions from unconventional speaker
Bicycle	Look at the thing with two-wheels, handlebars, pedals, and a seat.
Soap	Look at the thing that you use with water for washing your hands in the bathroom.
Balloon	Look at the round thing you blow air into that people have at parties.
Notebook	Look at the thing made of paper that you can write or draw in.
Candle	Look at the thing made of wax that you light and blow out.
Bottle	Look at the thing that you fill with milk and give to babies to drink from.
Cupcake	Look at the yummy dessert that is round and has icing on top.
Eraser	Look at the pink thing that is used to take away mistakes when you write with a pencil.
Grapes	Look at the round, green, juicy fruit that grows on vines.
Tree	Look at the plant that has a trunk, branches, and many leaves.
Plane	Look at the thing that flies in the sky and that takes people on long trips.
Glasses	Look at the things you put over your eyes to help you see better.
Chair	Look at the piece of furniture that you can sit on and eat dinner.
Squirrel	Look at the brown animal with a furry tail that lives in trees.
Carrot	Look at the orange vegetable that grows in the ground.
Bat	Look at the long, brown thing that you hit balls with

Figure S1

The proportion of fixations to matching (target and competitor objects) and mismatching (distractors and contrast object) items over time, across critical trials



Note. The dashed line reflects the offset of adjective.

2.7. References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
<https://doi.org/10.18637/jss.v067.i01>
- Berman, J. M., Graham, S. A., Callaway, D., & Chambers, C. G. (2013). Preschoolers use emotion in speech to learn new words. *Child development*, 84(5), 1791-1805.
<https://doi.org/10.1111/cdev.12074>
- Bohn, M., Tessler, M. H., Merrick, M., & Frank, M. C. (2021). How young children integrate information sources to infer the meaning of words. *Nature human behaviour*, 5(8), 1046-1054.
- Brown-Schmidt, S. (2009, Oct). The role of executive function in perspective taking during online language comprehension. *Psychon Bull Rev*, 16(5), 893-900.
<https://doi.org/10.3758/PBR.16.5.893>
- Brown-Schmidt, S., Gunlogson, C., & Tanenhaus, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition*, 107(3), 1122-1134. <https://doi.org/10.1016/j.cognition.2007.11.005>
- Cane, J. E., Ferguson, H. J., & Apperly, I. A. (2017). Using perspective to resolve reference: The impact of cognitive load and motivation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(4), 591. <https://doi.org/10.1037/xlm0000345>
- Davies, C., Lingwood, J., & Arunachalam, S. (2020). Adjective forms and functions in British English child-directed speech. *Journal of Child Language*, 47(1), 159-185.

- Davies, C., Lingwood, J., Ivanova, B., & Arunachalam, S. (2021). Three-year-olds' comprehension of contrastive and descriptive adjectives: Evidence for contrastive inference. *Cognition*, 212, 104707. <https://doi.org/10.1016/j.cognition.2021.104707>
- Dink, J., & Ferguson, B. (2018). *eyetrackingR*. In (Version 0.1.8) <http://www.eyetracking-R.com>
- Falkum, I. L., Recasens, M., & Clark, E. V. (2017). “The moustache sits down first”: on the acquisition of metonymy. *Journal of Child Language*, 44(1), 87-119. <https://doi.org/10.1017/S0305000915000720>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*, 39(2), 175-191. <https://doi.org/10.3758/bf03193146>
- Gardner, B., Dix, S., Lawrence, R., Morgan, C., Sullivan, A., & Kurumada, C. (2021). Online pragmatic interpretations of scalar adjectives are affected by perceived speaker reliability. *PloS one*, 16(2), e0245130. <https://doi.org/10.1371/journal.pone.0245130>
- Gazdar, G. (1979). A solution to the projection problem. In *Presupposition* (pp. 57-89). Brill.
- Graham, S. A., Stock, H., & Henderson, A. M. (2006). Nineteen-month-olds' understanding of the conventionality of object labels versus desires. *Infancy*, 9(3), 341-350. <https://doi.org/10.1371/journal.pone.0131215>
- Grice, H. P. (1975). Logic and conversation. In *Speech acts* (pp. 41-58). Brill.
- Grodner, D., & Sedivy, J. C. (2011). 10 The Effect of Speaker-Specific Information on Pragmatic Inferences. In *The processing and acquisition of reference* (Vol. 2327, pp. 239-272). MIT Press. <https://doi.org/10.7551/mitpress/9780262015127.003.0010>

- Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, 49(1), 43-61. [https://doi.org/10.1016/S0749-596X\(03\)00022-6](https://doi.org/10.1016/S0749-596X(03)00022-6)
- Heller, D., Grodner, D., & Tanenhaus, M. K. (2008, Sep). The role of perspective in identifying domains of reference. *Cognition*, 108(3), 831-836.
<https://doi.org/10.1016/j.cognition.2008.04.008>
- Henderson, A. M., Graham, S. A., & Schell, V. (2015). 24-month-olds' selective learning is not an all-or-none phenomenon. *PloS one*, 10(6), e0131215.
<https://doi.org/10.1371/journal.pone.0131215>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?
Behavioral and brain sciences, 33(2-3), 61-83.
- Horn, L. (1984). Towards a new taxonomy for pragmatic inference: Q-and R-based implicature.
Meaning, form and use in context.
- Huang, Y. T., & Snedeker, J. (2013). The use of lexical and referential cues in children's online interpretation of adjectives. *Developmental psychology*, 49(6), 1090.
<https://doi.org/10.1037/a0029477>
- Huetting, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta psychologica*, 137(2), 151-171.
- Khu, M., Chambers, C. G., & Graham, S. A. (2020). Preschoolers Flexibly Shift Between Speakers' Perspectives During Real-Time Language Comprehension. *Child development*, 91(3), e619-e634. <https://doi.org/10.1111/cdev.13270>

- Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child development*, 76(6), 1261-1277. <https://doi.org/10.1111/j.1467-8624.2005.00849.x>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82(1), 1-26. <https://doi.org/10.18637/jss.v082.i13>
- Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. MIT press.
- Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & psychophysics*, 53(4), 372-380. <https://doi.org/10.3758/BF03206780>
- Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton, Florida : CRC Press.
- Mozuraitis, M., Chambers, C. G., & Daneman, M. (2015). Privileged versus shared knowledge about object identity in real-time referential processing. *Cognition*, 142, 148-165. <https://doi.org/10.1016/j.cognition.2015.05.001>
- Nadig, A. S., & Sedivy, J. C. (2002, Jul). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychol Sci*, 13(4), 329-336. <https://doi.org/10.1111/j.0956-7976.2002.00460.x>
- Nilsen, E. S., & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive psychology*, 58(2), 220-249. <https://doi.org/10.11575/PRISM/1179>

- Orena, A. J., & White, K. S. (2015). I forget what that's called! Children's online processing of disfluencies depends on speaker knowledge. *Child development, 86*(6), 1701-1709.
<https://doi.org/10.1111/cdev.12421>
- Pouscoulous, N., & Tomasello, M. (2020). Early birds: Metaphor understanding in 3-year-olds. *Journal of Pragmatics, 156*, 160-167. <https://doi.org/10.1016/j.pragma.2019.05.021>
- R Core Team. (2020). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ryskin, R., Kurumada, C., & Brown-Schmidt, S. (2019, Aug). Information Integration in Modulation of Pragmatic Inferences During Online Language Comprehension. *Cogn Sci, 43*(8), e12769. <https://doi.org/10.1111/cogs.12769>
- Schulze, C., Grassmann, S., & Tomasello, M. (2013). 3-Year-old children make relevance inferences in indirect verbal communication. *Child development, 84*(6), 2079-2093.
<https://doi.org/10.1111/cdev.12093>
- Scofield, J., & Behrend, D. A. (2008). Learning words from reliable and unreliable speakers. *Cognitive Development, 23*(2), 278-290. <https://doi.org/10.1016/j.cogdev.2008.01.003>
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*(2), 109-147. [https://doi.org/10.1016/S0010-0277\(99\)00025-6](https://doi.org/10.1016/S0010-0277(99)00025-6)
- Stiller, A. J., Goodman, N. D., & Frank, M. C. (2015). Ad-hoc implicature in preschool children. *Language Learning and Development, 11*(2), 176-190.
<https://doi.org/10.1080/15475441.2014.927328>

- Syrett, K., LaTourrette, A., Ferguson, B., & Waxman, S. R. (2019). Crying helps, but being sad doesn't: Infants constrain nominal reference online using known verbs, but not known adjectives. *Cognition*, *193*, 104033. <https://doi.org/10.1016/j.cognition.2019.104033>
- Thacker, J. M., Chambers, C. G., & Graham, S. A. Preschoolers' Attention to Emotional Prosody: Pragmatic Adjustment in Response to Speaker Conventionality *Manuscript submitted for publication*.
- Thacker, J. M., Chambers, C. G., & Graham, S. A. (2018a). Five-year-olds' and adults' use of paralinguistic cues to overcome referential uncertainty. *Frontiers in psychology*, *9*, 143. <https://doi.org/10.3389/fpsyg.2018.00143>
- Thacker, J. M., Chambers, C. G., & Graham, S. A. (2018b). When it is apt to adapt: Flexible reasoning guides children's use of talker identity and disfluency cues. *Journal of experimental child psychology*, *167*, 314-327. <https://doi.org/10.1016/j.jecp.2017.11.008>
- Tribushinina, E., & Gillis, S. (2012). The acquisition of scalar structures: Production of adjectives and degree markers by Dutch-speaking children and their caregivers. *Linguistics*, *50*(2), 241-268.
- Trueswell, J. C. (2008). Using eye movements as a developmental measure within psycholinguistics. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 73-96). John Benjamins.

**CHAPTER THREE: Five-year-olds' sensitivity to knowledge discrepancies about object
identity during language comprehension**

Abstract

In everyday communication, children experience situations where their knowledge or perspectives differ from those of their communicative partner. The present study examined this issue in the context of real-time language comprehension, focusing on five-year-old children's ability to manage knowledge discrepancies about the identity of mutually visible objects. In Experiment 1, we examined five-year-olds' ability to manage privileged knowledge about an object's identity. Using a referential communication task, we tested children ($n = 60$) in either a shared-knowledge condition— where both the child and the speaker knew the identity of a visually misleading object (e.g., a candle that looked like an apple) — or a privileged-knowledge condition— where only the child knew the identity of the visually misleading object. Of interest was whether children could suppress private knowledge while processing a phonologically-related word (e.g., “Look at the *candy*”). Results showed children did not inhibit this knowledge during the early moments of referential interpretation. In Experiment 2 ($n = 30$), we contrasted the privileged knowledge condition in Experiment 1 with the more traditional scenario used to test common ground use, where the child knows the speaker cannot see certain display objects. Results confirmed a stronger ability to manage discrepancies in the latter case. Together, the findings demonstrate differences in children's ability to manage distinct types of knowledge discrepancies during real-time language comprehension.

Keywords: Perspective-taking, Real-time comprehension, Appearance-reality distinction, Physical co-presence, eye tracking

3.2. Introduction

People bring different beliefs, perspectives, and experiences to their communicative exchanges. For conversations to proceed effectively, conversational partners must distinguish and track knowledge that they share (i.e., information in common ground) and do not share (i.e., information in privileged ground). For example, a listener would not be confused by the request "Can you pass me the pen", if, in addition to a pen visible to them and the speaker, they knew of yet another pen inside their backpack. Previous studies have demonstrated that both adults and children use their assumptions about shared knowledge to guide both the production and comprehension of referential descriptions (e.g., Clark et al., 1983; Clark & Wilkes-Gibbs, 1986; Epley et al., 2004; Fussell & Krauss, 1989; Gibbs Jr et al., 1988; Matthews et al., 2006; Nadig & Sedivy, 2002; Nilsen & Graham, 2012; Wilkes-Gibbs & Clark, 1992 among others). To date, most developmental studies on this topic have focused on whether children monitor information that is visually shared or not shared with a conversational partner. In the present study, we explore this question from a complementary perspective, examining how 5-year-olds manage knowledge discrepancies about the *identity* of objects in real-time language comprehension.

As background, most studies examining conversational partners' attention to common ground (i.e., shared knowledge) in real-time language comprehension have used variants of the Visual World eye-tracking paradigm in which there are discrepancies in the visual perspectives of a listener and a speaker. These discrepancies often involve cases where one individual can see one or more objects that are not in the field of vision of the other because of a physical barrier or a visual cue signaling that a screen-displayed object is not present on the respective screen of their communicative partner (e.g., Barr, 2008; Barr & Keysar, 2002; Brown-Schmidt et al., 2008; Hanna & Tanenhaus, 2004; Heller et al., 2008 among others; Keysar et al., 2000; Keysar et al.,

2003). The efficiency of listeners' perspective-taking is marked by the pattern and timing of their eye movements as spoken instructions unfold (e.g., which objects they temporarily consider as potential referents and the speed with which they identify the intended target).

Early research using this paradigm reported an egocentric bias among adults and children whereby listeners showed a tendency to consider objects in privileged ground (visible to only themselves) as candidates for an unfolding description provided by the speaker (e.g., Barr, 2008; Barr & Keysar, 2002; Epley et al., 2004; Keysar et al., 2000; Keysar et al., 2003). A growing body of recent studies, however, has demonstrated that both adults and children can effectively draw on knowledge of which referential candidates are mutually visible to guide real-time referential interpretation (e.g., Brown-Schmidt & Heller, 2014; Ferguson & Breheny, 2012; Hanna & Tanenhaus, 2004; Hanna et al., 2003; Heller et al., 2008; Heller et al., 2016; Khu et al., 2020; Mozuraitis et al., 2015; Nilsen & Graham, 2009; Saryazdi & Chambers, 2021). Further, there is evidence that a number of factors, including the cognitive demands of a given task, or the degree of linguistic match between a target description and a competitor object, can affect the level of success that listeners demonstrate in managing their privileged knowledge (see Brown-Schmidt & Heller, 2018; San Juan et al., 2015 for reviews and discussion).

Despite the methodological focus on cases involving shared versus distinct visual perspectives, everyday communication reflects many other ways in which individuals come to understand and track shared and privileged knowledge, which in turn may be more challenging for listeners. For example, taking another person's *emotional* perspective (when it conflicts with one's own emotional response to a situation), might be more challenging for children than managing discrepancies in line-of-sight visual perspective (Wellman, 2018). In support of this, Khu et al. (2018) report processing delays in preschool children's use of emotional prosody to

infer a speaker's emotional perspective. This stands in contrast to the rapid real-time effects of emotional prosody found in non-conversational tasks that do not require perspective-taking (e.g., Berman et al., 2010; Berman, Graham, & Chambers, 2013).

In the present studies, we examine another type of perspective-taking in the context of real-time language processing, namely, how effectively child listeners integrate a speaker's knowledge about the identity of an object when this knowledge conflicts with their own. This question is motivated by a rich literature in developmental psychology that documents the emergence of children's sensitivity to the distinction between the *appearance* of objects and their *reality*. In a typical appearance-reality task, children are presented with objects whose visual properties are at odds with their actual conceptual category (e.g., a sponge that looks like a rock). After being shown or told the real identity of objects, children are asked questions related to appearance (i.e., "what does it look like?") and actual identity (i.e., "what is this really?"). Studies have shown that, between 4 and 5 years of age, children can begin to explicitly identify the identity of a deceptive object (e.g., Flavell, 1993; Flavell et al., 1983; Gopnik & Astington, 1988) and appreciate that another person will misidentify that object based on its appearance (Apperly & Robinson, 1998; Gopnik & Astington, 1988). Further, when instructions and task demands are simplified, even 3-year-old children successfully distinguish between the appearance and reality of objects (e.g., Hansen & Markman, 2005; Moll & Tomasello, 2012; Sapp et al., 2000). For example, when children are asked to choose an object that only "looks like chocolate", they are more likely to point to the eraser that resembles a bar of chocolate than a real chocolate bar (Moll & Tomasello, 2012). Understanding the distinction between appearance and reality and the relation to another's knowledge is highly relevant, as failure to do

so can entail misconceptions or misunderstandings that have consequences for social interactions (Flavell et al., 1983; Moll & Tomasello, 2012).

Given that young children can distinguish between the appearance and reality of objects in offline tasks, our goal is to explore whether and how child listeners manage knowledge discrepancies of this type in real-time communication. We predict children may experience more difficulty with this type of knowledge mismatch than the more well-studied case involving the difference in *visual* perspective. Not only do appearance/reality cases involve more distinctly cognitive types of mentalizing, but it has also been suggested that managing knowledge mismatch of this type is inherently more demanding, as it requires one to represent the way others represent an object, as opposed to simply tracking others' beliefs or visual access (Butterfill & Apperly, 2013). Thus, even though young children can demonstrate an explicit understanding of another's false belief about object identity, it is unclear whether they will readily use this type of information to guide moment-by-moment language processing. Indeed, studies of syntactic processing have shown that preschool children's offline understanding of certain conventions of language use is not used to guide and revise their online interpretation (Hurewitz et al., 2000). An understanding of how children deal with appearance/reality discrepancies in real-time will broaden our understanding of children's perspective-taking abilities beyond the more traditionally studied contexts involving visual perspective.

In the first experiment, we examined 5-year-old children's sensitivity to privileged versus shared knowledge about the identity of mutually visible objects during online comprehension. We used "visually misleading" objects to create a knowledge discrepancy between the child listener and a speaker, such as a candle that looks like an apple. We monitored eye movements as children followed instructions produced by a speaker who was either knowledgeable about the

actual identity (shared knowledge condition) or ignorant of the actual identity (privileged knowledge condition) of visually misleading objects. On critical trials, a visually misleading object (e.g., candle) served as a competitor object for a target object named in the instruction (e.g., candy) due to shared initial sounds. Typically, this phonological overlap creates temporary ambiguity regarding the upcoming word, leading listeners to momentarily fixate to a competitor object en route to identifying the target item. In the present experiment, however, the potential for the temporary ambiguity depended upon children's consideration of the speaker's knowledge about object identity. That is, if the ignorant speaker (privileged knowledge condition) was to refer to the visually misleading object, it should be on the basis of its appearance (e.g., apple) for which the corresponding label does not share onset sounds with the target (candy). It is worth noting that temporary ambiguity within a single word provides a comparatively stringent test of listeners' use of perspective information during real-time language processing compared to cases involving a description that is ambiguous in its entirety (see also Barr, 2008). However, this approach helps ensure that the full-blown linguistic ambiguity that would occur otherwise does not lead listeners to strategically attend to perspective cues. Thus, if children show little or no competitor fixations as the target word unfolds when the speaker is unaware of the actual identity of the deceptive object, this outcome will provide strong evidence for children's effective use of the speaker's perspective in online comprehension.

3.3. Experiment 1

In Experiment 1, we investigated 5-year-olds' real-time communicative perspective-taking in situations where the speaker either did or did not know the identity of a visually misleading object. We adapted a version of the visual world paradigm to create situations wherein a child listener and an adult speaker either shared knowledge (shared knowledge

condition) or did not share knowledge (privileged knowledge condition) about the actual identities of visually misleading objects. In the latter case, only the child listener was aware of the true identity of misleading objects. On critical trials, a visually misleading object (e.g., a candle that looked like an apple) served as a phonological competitor for the target word in an instruction produced by the adult speaker (e.g., *Look at the candy*). However, as noted earlier, the potential for this object to function as a competitor should in principle depend on whether knowledge of true object identity was shared. That is, from child listener's perspective, when both the child listener and the adult speaker know the identity of the visually misleading object (i.e., candle), its presumed label will have sound overlap with the target word in the spoken instruction (*candy*). Thus, in the shared knowledge condition, child listeners should consider the visually misleading object as the target word unfolds. In contrast, when the speaker is unaware of the true identity of the misleading object (privileged knowledge condition), the competitor object should not, in principle, be considered because the speaker's label for that object does not overlap with the target. If, however, children do not effectively manage their privileged knowledge about the competitor object's identity, they will show competitor fixations in both the shared and privileged knowledge conditions.

3.3.1. Method

3.3.1.1. Participants

Data from 60 5-year-old children (30 boys; $M_{age} = 5.73$ years, $SD = .14$, range = 5 years 6 months to 5 years 11 months) were included in the final sample. Children were randomly assigned to one of the two knowledge state groups ($n = 31$ in shared knowledge; $n = 29$ in privileged knowledge), which are described in more detail below. All children spoke English as their primary language (greater than 60% of the time). The current sample size is sufficient to

detect a medium effect size ($f = 0.25$) with 95% power in a conventional repeated-measures analysis of variance (ANOVA) for a relevant interaction with alpha at .05 (Faul et al., 2007). An additional 23 children were tested but excluded from the final sample for the following reasons: difficulties understanding task manipulations/instructions ($n = 1$), technical error ($n = 4$), experimenter error ($n = 2$), unsuccessful calibration ($n = 2$), failure to meet language criteria ($n = 1$), parental report of developmental disorders or delay ($n = 4$), one twin whose sibling had already participated, and insufficient eye-gaze data collected ($n = 8$). Children were from diverse ethnic backgrounds (Caucasian, 70%; Multiethnic, 17%; East Asian, 5%; Hispanic/Latino, 3%; South Asian, 3%; Middle Eastern, 2%), and the majority were from families where parents had at least some post-secondary education (84%). This study was approved by the [removed for review] under the project name Communicative Development (removed for review).

3.3.1.2. Apparatus

The visual stimuli were displayed on a 46-inch screen. The experiment was programmed and conducted using E-Prime software with Tobii extensions. Children's eye movements were recorded using a Tobii pro x3-120 eye tracker. A five-point gaze calibration was administered using Tobii ClearView software. Gaze calibration was considered as accurate if at least three dots were recognized by the eye tracker. Areas of interest (AOIs) were specified before testing, based on the location of each object on the display. Children's gaze position was logged every 8.3 ms and a fixation was registered if gaze to a given location lasted longer than 95 ms. Auditory stimuli were played through speakers located either side of the screen.

3.3.1.3. Materials

A total of 13 visually misleading objects were used. The appearance of the visually misleading objects differed from their actual identity (e.g., a crayon that looks like a block).

These objects were carefully chosen so that the object categories for both their appearance and identity would be familiar to children. Of the 13 objects, one object was used in a practice trial, three objects were presented in filler trials, and the remaining nine objects were used in critical trials.

Table 1

List of target objects and three types of competitors

Target	Regular phonological competitor	Knowledge-based phonological competitor (looks like)	Visual control object
candy	candle	candle (apple)	apple
chalk	chocolate	chocolate (coin)	coin
pennies	pen	pen (flower)	flower
soldier	soap	soap (duck)	duck
coat	comb	comb (cookie)	cookie
yogurt	yo-yo	yo-yo (baseball)	baseball
flag	flashlight	flashlight (pig)	pig
bottle	box	box (lego)	lego

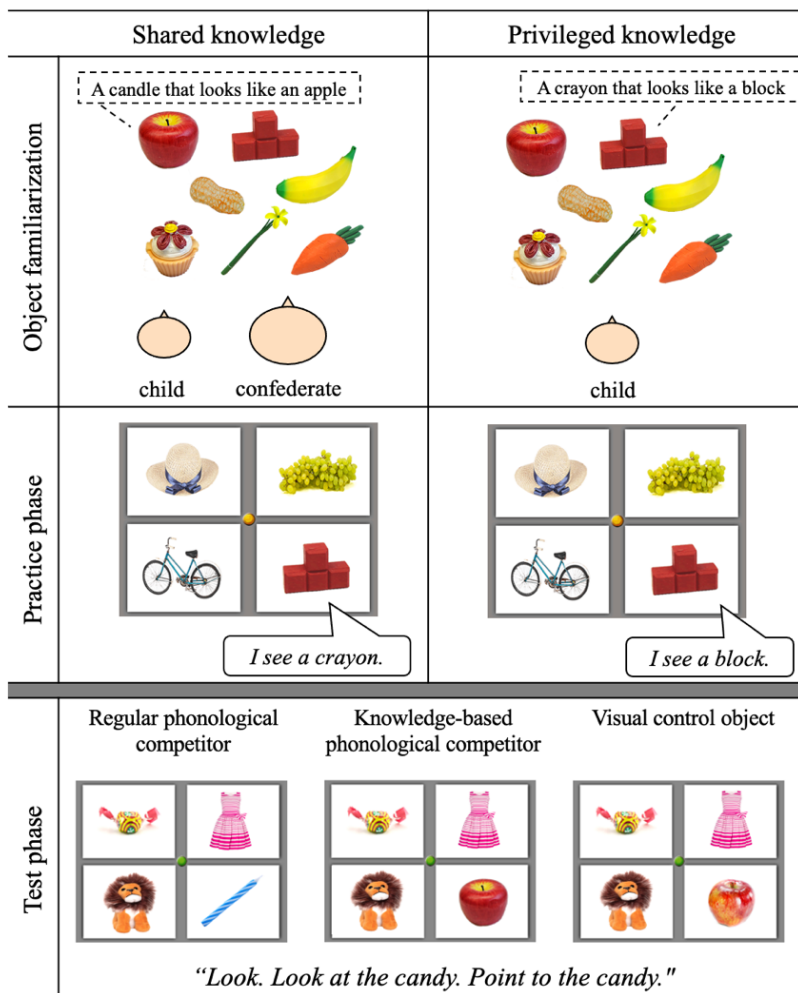
Note. A regular candle (A), candle that looks like an apple (B), and regular apple (C), illustrating different types of competitors.

The task consisted of a practice phase (2 trials) and a test phase (18 trials). The visual display for each trial consisted of an array of images constituting a pre-instruction display and an instruction display, presented sequentially. In the pre-instruction display, four cards with question marks were presented. These cards indicated the location of each object, one in each corner of the screen, and lasted about 1 second. In the instruction display, the question marks were replaced with four objects (see Figure 1 for example displays): one target object, one competitor object, and two unrelated items (based on the spoken instructions). On critical trials,

we systematically varied the type of competitor occurring in a given instruction display (see Table 1 for a list of items). The *regular phonological competitor* was an object whose name shared onset sounds with the target name. For example, the regular phonological competitor for the target *candle* was a candy, and a chocolate for the target *chalk*. For the *knowledge-based phonological competitor*, a corresponding visually-misleading object was used. This object's appearance was misleading in relation to its actual identity, and only the label corresponding to this true identity shared onset sounds with the target object (e.g., a candle that looks like an apple for the target noun *candy*, a chocolate that looks like a coin for the target noun *chalk*). This competitor, thus, had potential to function as a phonological competitor for the target object depending on whether the speaker possessed knowledge of its true identity. Lastly, the *visual-control object* was an actual instance of the category corresponding to the knowledge-based competitor's appearance (e.g., an apple, a coin).

Figure 1

Testing sequence with examples of visual and auditory stimuli.



Note. A regular candle, candle that looks like an apple, and regular apple, illustrating different types of competitors. The images are recreated for illustrative purposes only. The visual displays used in the studies are similar but not identical to this figure.

The target and competitor objects always appeared on the opposite side of the display. For example, if the target object was in the top-left corner, the competitor was in the bottom-right corner. The location of target object was counterbalanced across trials. Each array

presented in critical trials was fixed except the type of competitor changed depending on the experimental condition. Three test orders were created to cycle the pairing of arrays to conditions. Across participants, all three types of competitors occurred in each array (see Figure 1). However, a participant saw a given array only once.

The nine filler trials consisted of three different types: three trials involving a phonological competitor (e.g., a sandwich as target, a sandal as competitor, a doll and a dinosaur for unrelated objects), three trials involving a visually misleading object whose name had no phonological overlap with the target name (e.g., a cat as target, an eraser that looks like a peanut for the visually misleading object, a toothbrush and an egg for unrelated objects), and the other filler trials involving no object which had the same onset sounds with the target or a deceptive appearance (e.g., a book as target, a book, a phone, and a train for unrelated objects). Filler and critical trials were presented in a pseudorandomized order where the same type of trial did not occur more than two consecutive times.

Auditory stimuli were pre-recorded using a child-directed speech by a female native English speaker. The speaker recorded utterances for the practice phase as well as the critical instructions in the test phase. The critical instructions were presented after 500 ms of silence. (“Look. Look at the [target object]. Point to the [target object]”). They were presented in conjunction with the visual displays, and thus seemed to be spontaneously produced by the adult confederate from another room.

3.3.1.4. Procedure

Children were randomly assigned to either the shared knowledge condition or privileged knowledge condition. Each child was first brought to a control room located adjacent to the testing room and met the confederate who was sitting in front of a computer while wearing a

headset microphone. The child was told that the confederate would be playing the same game as the child and was able to hear and talk through the headset. The confederate did not say anything but smiled and waved to the child. The purpose of this setup was to suggest to the child that they were interacting with the confederate in real-time and thus to encourage the child to be engaged during the task. As noted above, all audio stimuli in fact were pre-recorded to ensure consistency.

After meeting the confederate in a control room, the procedure in the shared knowledge condition was such that both the child and the confederate were brought into the testing room. The child was seated across a small square table from the experimenter, and the confederate sat diagonally between the child and the experimenter. In the privileged knowledge condition, the confederate stayed in the control room and only the child entered the testing room.

Object familiarization phase. Before starting the communication task, the child participant and the confederate (shared knowledge condition) or only the child participant (privileged knowledge condition) were familiarized with the visually misleading objects that would be shown in the communication task. The experimenter first sequentially presented seven visually misleading objects (one used on a practice trial, three on filler trials, and the other on critical trials) and asked the child to name each one (e.g., “Look at this. What is it?”). We expected that children would answer based on the appearance of the object. If the child was hesitant to answer, then the experimenter prompted the child to say what the object looked like. The experimenter then explained the actual identity of the object with its function (e.g., “It looks like an apple, but this really is a candle. You could light it and blow it out”) and handed it to the child to look at and touch for 5 seconds. The same procedure was repeated for all the visually misleading objects used on critical trials. After the visually misleading objects had been presented, the experimenter presented images of those objects on the screen and again asked the

child what each object was. If the child gave an answer based on the object's appearance, the experimenter prompted them to say what it really was. Once the child correctly identified the actual identity of the objects on the screen, the experimenter asked the child what a new person would think these objects are (e.g., "Remember, I have never shown these objects to anyone before. If one of my friends came in and I asked her what it is, what do you think she would say?"). If the child provided the actual identity of the object, the experimenter corrected their response (e.g., "Remember, they do not know it is a candle. They would say it is an apple."). In the shared knowledge condition, the confederate left the testing room once the child completed this false belief understanding task. The object familiarization procedure was identical in the privileged knowledge condition with two exceptions: the confederate was not present in the testing room, and the experimenter emphasized that only the child would know the true identity of the visually misleading objects. One child in privileged knowledge condition was excluded from the analysis due to failure to understand another person's false beliefs about object identity.

Practice phase. After calibration of the eye tracker was completed, the experimenter told the child that the confederate could hear through the speakers from the other room, and demonstrated that when the speakers were turned off, the confederate cannot hear what the experimenter was saying. The practice trials then began. The first practice trial presented four familiar objects. For each object, the experimenter pretended to ask the confederate what she was looking at, and played the pre-recorded response, which named each object. After naming all objects on the display, the experimenter asked the child to listen carefully and follow the instructions from the confederate ("Look! Look at the ring. Point to the ring").

The second practice trial showed three familiar objects and one of the visually misleading objects presented in the object familiarization phase (e.g., a crayon that looks like a block). The

experimenter again pretended to interact with the confederate and asked them to name each object. The labelling of the visually misleading object from the confederate, however, varied depending on the condition. In the shared knowledge condition, the visually misleading object was named as its actual identity (e.g., “I see a crayon.”). The experimenter asked whether she was sure about it, and the pre-recorded voice confirmed it. Then the experimenter turned off the speaker and asked the child whether the confederate knew what it really was. In the privileged knowledge condition, the visually misleading object was named based on its appearance (e.g., “I see a block.”). The experimenter and the confederate speaker followed the same conversational dialogue as in the shared knowledge condition. All children included in the final sample correctly answered the question querying the confederate’s knowledge about the visually misleading object.

Test phase. There were 18 trials in total, which consisted of nine critical trials and nine filler trials. On each trial, the experimenter named all objects while the microphone and speaker used to communicate with the confederate were turned off (e.g., “Here we have a candy, a dress, a lion, and a candle”). After labeling, the experimenter turned on the speaker and pretended to let the confederate know that the child was ready and played the instruction. The instruction was of the form "*Look! Look at the [object name]! Point to the [object name]!*" across trials. Once the communication game was done, the experimenter took off a headset and confirmed the knowledge state manipulation by asking the child about the confederate’s knowledge state (i.e., whether the confederate knew the true identity of the visually misleading objects).

3.3.2. Data Analyses

The measures of interest involve eye fixations during the unfolding noun (an 800 ms interval beginning 200 ms after noun onset and ending 1000 ms after noun onset; average noun

duration was 817 ms). The 200 ms margins were added to account for the time required to program and execute eye movements (Matin, Shao, & Boff, 1993). Fixations launched before the 200 ms onset of the analysis interval were excluded to ensure that the data patterns were not influenced by any lingering fixations launched before the noun was processed.

Prior to analyses, we excluded trials in which track loss occurred more than 70% of samples within the noun interval, using the `eyetrackingR` package (Dink & Ferguson, 2018). Participants for whom all three trials were excluded for a given experimental condition were removed from the analysis due to insufficient eye gaze data. For the remaining participants, the track loss criterion resulted in 76 of the original 540 trials being excluded (14.05% of the original data set).

Statistical analyses were performed using R version 4.0.3 (R Core Team, 2020). For these analyses, we aggregated fixations into 50 ms time bins and calculated the proportion of fixations to the target and competitor by dividing the sum of all fixations to the displayed objects by the fixations to each target and competitor. Target advantage scores were then calculated by subtracting the proportion of fixations to the competitor from the proportion of fixations to the target, thereby providing a single value reflecting the extent to which participants considered the target relative to the competitor. These proportions were rescaled to range between 0 and 1 and were then transformed using a logit transformation to enable analyses using conventional linear methods.

The linear mixed effect analyses were performed using the ‘`lme4`’ package (Bates et al., 2015) as well as the ‘`lmerTest`’ package (Kuznetsova et al., 2017) in R. Knowledge state (shared/privileged) and competitor type (regular/knowledge-based/visual control), along with their interaction, were included as fixed effects. Our models were based on a maximal random

effects structure and included random intercepts for participants and items as well as a by-participant random slope for competitor type, and by-item random slopes for knowledge state and competitor type, and their interaction.

In the model, knowledge state was sum coded (1 = shared knowledge vs. -1 = privileged knowledge). For competitor type, we used a linear contrast and a quadratic contrast for the analysis, with the goal of assessing which provided a better fit for the data. The logic in defining these patterns is based on the approach adopted in Heller et al. 2016, where the visual and linguistic factors that may drive fixations from a default egocentric perspective are isolated separately from possible effects of perspective-taking. The linear pattern (-1 = knowledge-based phonological competitor, 0 = regular phonological competitor, 1 = visual object) corresponds to a scenario where target advantage scores increase in a steady manner across the levels of the competitor manipulation (knowledge-based phonological competitor, regular phonological competitor, visual control object). This pattern reflects an assumption that the knowledge-based phonological competitor will strongly draw children's attention, as it would be the most attention-grabbing object for children (due to the novelty/inherent interest linked to visually deceptive objects; see Berman, Graham, Callaway, et al. (2013) for a finding about children's tendency to look at a visually complex object) and its label shares the onset sounds with the target object from the listener's perspective. The target advantage scores in the regular phonological competitor condition should be higher as the competitor has less of a novelty value even though there is still phonological overlap. Lastly, in the visual control object condition, the lack of phonological overlap with the target should entail few looks to the competitor, in turn further boosting the target advantage score. The quadratic version of the competitor type manipulation (0.5 = knowledge-based phonological competitor, -1 = regular phonological

competitor, 0.5 = visual object) will detect changes in the linear trend such as a decrease in the target advantage scores in a regular phonological competitor. It would reflect the effect of object appearance such that the target advantage scores will be lower in the knowledge-based competitor and the visual control object compared to the regular phonological competitor because the appearances of the first two types of competitors do not share the phonological sounds with the target object.

If children consider a speaker's knowledge in the privileged knowledge condition, they will treat the knowledge-based phonological competitor based on its appearance, similar to the visual control object. On the other hand, children in the shared knowledge condition will show an overall increase in target advantage scores across the knowledge-based phonological competitor, the regular phonological competitor, and the visual control object. However, if children consider their privileged knowledge, then their visual consideration of objects on display would be the same as those in the shared knowledge condition. Thus, the linear trend over competitor types will be shown in both knowledge states.

3.3.3. Results and Discussion

The average target advantage scores across conditions are shown in Figure 2. Scores above 0 indicate a greater tendency to look at the target object as the target noun unfolds, and scores below 0 indicate a greater tendency to look at the competitor object. Visual inspection of the patterns shows a successive increase in the scores across the levels of the competitor type manipulation. Across both levels of the knowledge state manipulation, there was a strong tendency to fixate knowledge-based phonological competitors, a weaker tendency with regular phonological competitors (entailing roughly equal consideration of the target and competitor), and a stronger tendency to fixate the target object when the competitor was replaced with a

visual control object. Importantly, however, there does not appear to be an effect of knowledge state (shared vs. privileged knowledge about the identity of competitor items).

The summary of the statistical analyses is provided in Table 2. Confirming the apparent pattern observed in Figure 2, the results indicated a linear increase in target advantage scores across the competitor type manipulation ($\beta = 0.28$, $SE = 0.09$, $t = 3.18$, $p = .008$). The quadratic pattern of competitor type and the main effect of knowledge state were not significant. Critically, there was no interaction of knowledge state with either kind of competitor type effect ($ps > .05$). Thus, children's consideration of the objects did not differ significantly depending on their knowledge about the speaker's awareness of the true identity of the visually misleading objects.

Figure 2

Average target advantage scores within the noun interval, Experiment 1

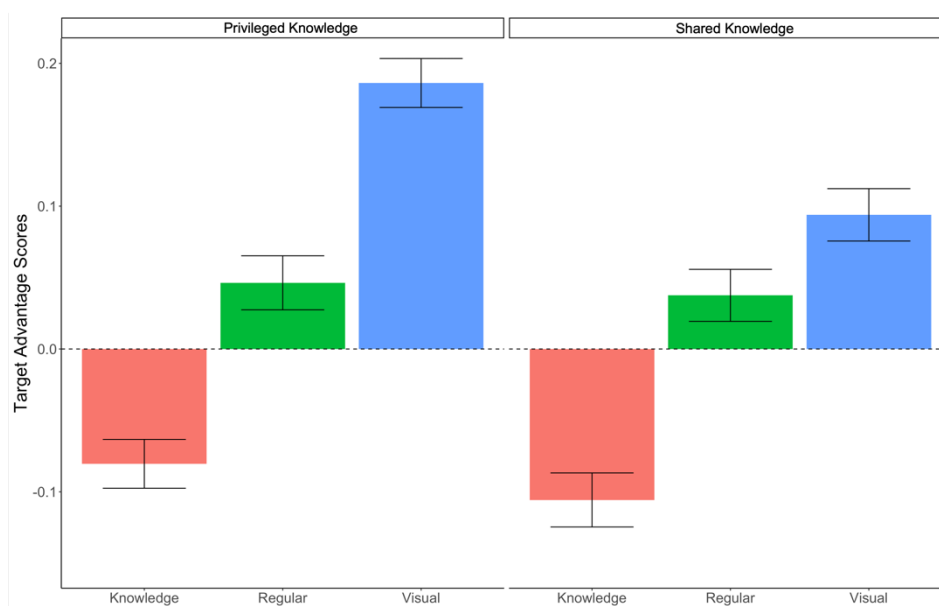


Table 2*Summary of statistical results, Experiment 1.*

	Estimate	<i>SE</i>	<i>t</i>	<i>p</i>
(Intercept)	0.09	0.10	0.91	.394
Knowledge state	0.09	0.07	1.26	.243
Linear	0.28	0.09	3.18	.008
Quadratic	0.01	0.12	0.08	.935
Knowledge state x Linear	0.02	0.08	0.30	.764
Knowledge state x Quadratic	0.06	0.14	0.44	.672

Results from Experiment 1 indicate that, as the target noun unfolded, children considered the actual identity of visually misleading objects (i.e., the knowledge-based phonological competitor) regardless of the speaker's knowledge state. The results also suggest that the early moments of comprehension in children are not tied to superficial types of perceptual information. This is because the outward appearance of this candidate is that of an object whose label bears relation to the target name. The salient visual information was therefore trumped by child listeners' internal representations of objects' conceptual categories. However, the finding that children did not reduce fixations to visually-misleading competitors in the privileged knowledge condition suggests important limitations in real-time perspective-taking. Specifically, this result shows that children had difficulty drawing on and applying knowledge that the speaker was naïve about the true identity of deceptive objects. This was the case even though the procedure made this lack of knowledge very apparent. This stands in contrast with the pattern found in studies exploring whether children draw on knowledge that speakers are entirely unaware of certain objects in the display, where effective use of common ground is in fact evident in the early moments of processing (Khu et al., 2020; Nadig & Sedivy, 2002; Nilsen & Graham, 2009).

This difference suggests that experimental methodologies involving occluded objects provide an incomplete picture of real-time perspective reasoning during language processing.

To confirm this idea in the context of a single study, we conducted a second experiment implementing the familiar occluded-objects manipulation but preserving the same materials and competitor type manipulations in Experiment 1. If children are more effective at drawing on common ground in this latter case, this would provide within-study evidence for the pattern of relative differences in perspective-taking scenarios involving object identify versus object co-presence. The results of Experiment 2 will also help a different explanation for children's difficulty in managing privileged knowledge about object identity. That is, it is possible that children's intrinsic interest in the novelty items we used as visually-deceptive objects steers the allocation of attention in a way that washes out effects of perspective-taking. Although we tried to reduce this concern by following a common practice whereby only those fixations beginning 200 ms after target noun onset were included (so that the gaze patterns used in the analysis were most likely driven by the linguistic signal), the attraction may nonetheless be very strong. If so, then children should continue to show a bias to fixate a visually-deceptive object even when they recognize the speaker cannot see it.

3.4. Experiment 2

Experiment 2 explored children's use of common ground using the same materials as Experiment 1 but adapted the scenario such that the child was aware that the speaker could not see the competitor object. To implement this with our screen-based displays, we introduced a "secret card" methodology to establish the speaker's ignorance of certain objects visible to the child participant. The secret card had a different background colour (blue) from other cards

(white), which indicated that objects on the blue card were visible only to the child participant, but not to the speaker.

3.4.1. Method

3.4.1.1. Participants

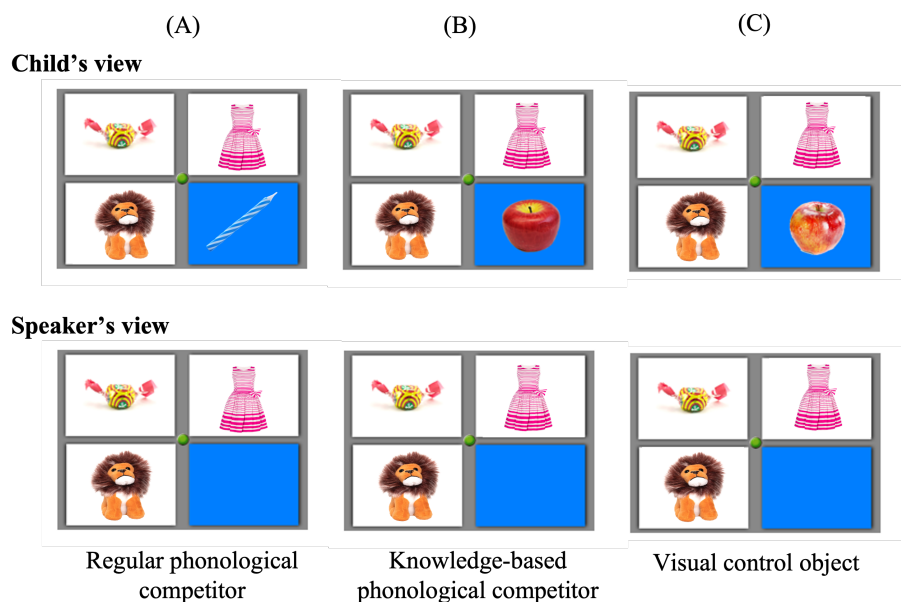
Data from 30 5-year-old children (15 boys; $M_{age} = 5.69$ years, $SD = .14$, range = 5 years 6 months to 5 years 11 months) were included in the analysis. All children used English as their primary language (greater than 60% of the time). This sample size is sufficient to detect a medium effect size ($f = 0.25$) with 95% power. An additional 8 children were tested but excluded from the final sample for the following reasons: difficulties understanding task manipulations/instructions ($n = 1$), technical error ($n = 2$), experimenter error ($n = 1$), unsuccessful calibration ($n = 1$), failure to meet language criteria ($n = 2$), and insufficient eye-gaze data ($n = 1$). Children were from diverse ethnic backgrounds (Caucasian, 67%; Multiethnic, 23%; Hispanic/Latino, 3%; South Asian, 3%; Middle Eastern, 3%), and the majority had parents who had completed at least some post-secondary education (85%).

3.4.1.2. Apparatus and Materials

The apparatus was identical to that used in Experiment 1. The stimuli were also the same with one exception. All nine critical trials and three filler trials (one in each type of filler trials) involved displays containing one object on the secret card visible to the child listener but not to the speaker (see Figure 3 for an example display). On critical trials, all competitor objects appeared on the secret card. In the filler trials involving a phonological competitor, the competitor again appeared on the secret card. In the other two filler trials containing a secret card, the secret card showed an unrelated object (e.g., an umbrella when the target was a towel).

Figure 3

Example of critical display in Experiment 2 depending on the type of competitor



Note. The competitor (bottom right) is (from left to right) either a regular candle, a candle that looks like an apple, or a regular apple. The blue background indicates a “secret card” where only the child can see the object depicted (top row) and the speaker sees only the blue card (bottom row). The images are recreated for illustrative purposes only. The visual displays used in the studies are similar but not identical to this figure.

3.4.1.3. Procedure

The overall procedure was similar to the shared knowledge condition in Experiment 1. After meeting the confederate, the child entered the testing room with her. The child and the confederate were introduced to the true identity of the visually misleading objects. Once the child answered the questions of another’s false beliefs about the visually misleading objects, the confederate left the testing room.

The first practice trial was the same as in Experiment 1. However, in the present experiment, the second practice trial differed. Although it involved three standard objects and one visually misleading object (i.e., a crayon that looks like a block), the card on which the

visually misleading object was displayed was blue. When the experimenter asked the confederate what she saw on the display, the (pre-recorded) voice answered that the blue card did not flip on her screen. This blue card was introduced to the child as a secret card, for which only the experimenter and the child could see the object depicted on the other side. Then the experimenter engaged in a brief guessing game about the object on the card with the confederate (again using pre-recorded utterances), where the confederate failed to guess correctly and expressed that it is hard to guess when she cannot see what is on that card. After this trial, the child was asked whether the confederate knew what was on the blue card. One child was removed from the final sample because their answer showed a failure to understand the confederate's limited visual access to objects on the blue card. The testing phase began after the practice trials. The procedure for the testing phase was identical to that of Experiment 1.

3.4.2. Data Analyses

Prior to analyses, we excluded trials in which track loss occurred in more than 70% of the noun interval. One participant who, as a result of this criterion, was left with no observations in one of the experimental conditions, was removed entirely. For the remaining participants, the track loss criterion resulted in 33 exclusions out of the original 237 trials (14.22% of the original data set).

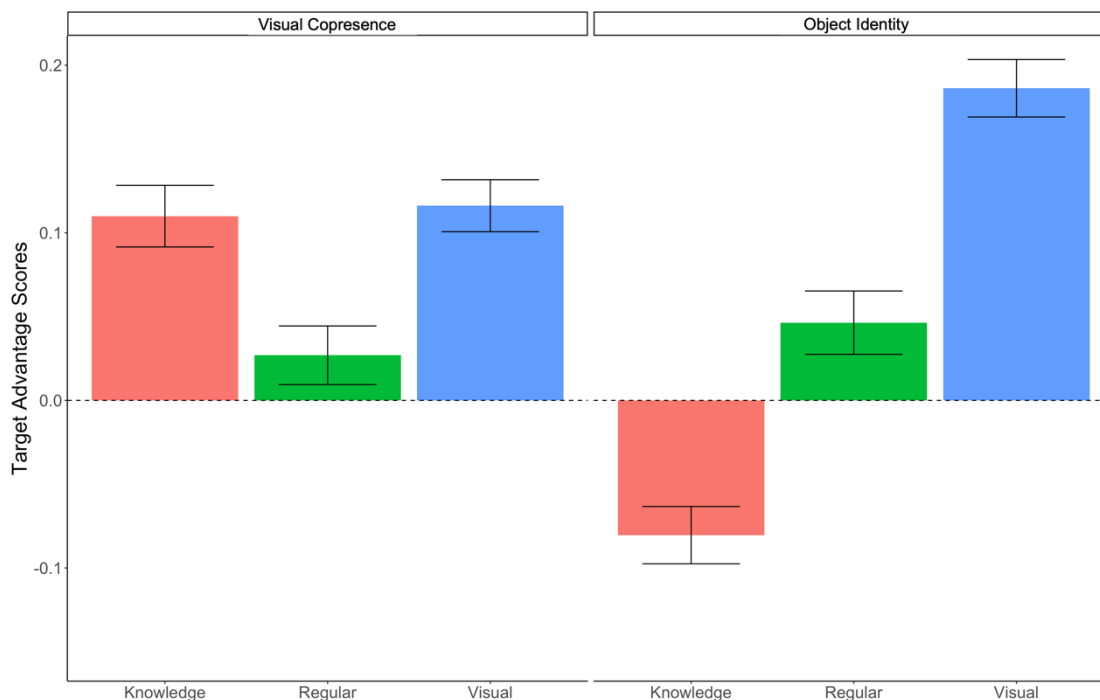
To compare children's ability to manage privileged knowledge about the contextual presence of objects versus objects' identity, we contrasted the results of the present experiment against the privileged knowledge condition from Experiment 1. Data preparation and the overall structure of the statistical model was the same as before.

3.4.3. Results and Discussion

We again calculated target advantage scores during the 800 ms noun interval. The left panel of Figure 4 shows the data from the present experiment, where the child listener had privileged knowledge about the presence of the competitor object. In this case, the pattern shows notably higher target advantage scores on average, reflecting reduced consideration of the competitor as the target noun unfolded. This contrasts with the corresponding condition from Experiment 2 (repeated as the right panel), involving privileged knowledge of object identity. As explained before, this latter pattern illustrates children's difficulty in suppressing privileged knowledge about object identity.

Figure 4

Average target advantage scores within the noun interval, Experiment 2



We included these two privileged knowledge types (visual copresence vs. object identity) in the statistical analysis alongside the manipulation involving competitor type. The statistical summary is provided in Table 3. There was a significant interaction between knowledge type and

the linear version of the competitor type contrast ($\beta = -0.15$, $SE = 0.06$, $t = -2.34$, $p = .020$). The other effects were not significant ($ps > .05$). Follow-up analyses were conducted to explore the interaction further. Reflecting the earlier analysis, the pattern in the privileged identity condition from Experiment 1 showed a significant linear trend whereby target advantage scores increased across competitor conditions ($\beta = 0.31$, $SE = 0.10$, $t = 3.23$, $p = .001$). However, there was no statistical difference across competitor conditions when the child knew the speaker could not see the competitor object ($p = .923$).

Table 3

Summary of statistical results, Experiment 2

	Estimate	SE	t	p
Main analysis				
(Intercept)	0.16	0.11	1.49	.174
Knowledge type	-0.01	0.06	-0.22	.829
Linear	0.16	0.08	2.04	.066
Quadratic	0.08	0.12	0.64	.536
Knowledge type x Linear	-0.15	0.06	-2.34	.020
Knowledge type x Quadratic	0.01	0.09	0.10	.922
Follow-up analyses				
<i>Physical co-presence</i>				
(Intercept)	0.15	0.14	1.03	.332
Linear	0.01	0.10	0.10	.923
<i>Object identity</i>				
(Intercept)	0.17	0.10	1.67	.131
Linear	0.31	0.10	3.23	.001

In summary, the results show that, when children were aware that their partner could not see the competitor object, their identification of the target noun was comparatively unaffected by

this object, regardless of competitor type. This indicates that, in this scenario, children were sensitive to the discrepancy between their own versus the speaker's knowledge about the available candidate for reference and used this knowledge to guide the early moments of processing. This finding is consistent with previous research showing that preschool children rapidly use their communicative partner's limited visual access to guide referential interpretation (Khu et al., 2020; Nadig & Sedivy, 2002; Nilsen & Graham, 2009). Importantly, this finding contrasts with the scenario where children's privileged information involves knowledge about the true identity of deceptive objects, where the competitor attracted substantially more attention. The fact that this pattern occurs only when there is shared visual access to the competitor indicates that children's apparent lack of perspective-taking in Experiment 1 cannot be simply due to an independent bias to fixate to visually misleading objects.

3.5. General Discussion

The present studies make several contributions to our understanding of the development of perspective-taking in real-time language processing during the preschool years. To begin, our results replicate and extend previous research documenting children's rapid sensitivity to others' *visual* perspectives in real-time language processing (Khu et al., 2020; Nadig & Sedivy, 2002; Nilsen & Graham, 2009). In Experiment 2, 5-year-olds' understanding that the speaker could not see a competitor led them to reduce consideration of that object during referential processing. This finding is impressive given that the visual access manipulation used in the present study is likely to be challenging for children. That is, because children were interacting with a remote speaker, they could not track directly the speaker's line of sight, nor was there a physical occluder to signal that an object was obscured from the speaker's perspective. Instead, children had to remember that a blue card meant that the object on the card did not appear on the

speaker's screen, infer the speaker's knowledge state based on that visual cue, and integrate this perspective information during language processing. This finding thus extends prior research by demonstrating that children consider a speaker's visual perspective in a remote communicative setting where they have to remember a speaker's knowledge state using a visual cue.

The results also inform our key question, namely whether children effectively manage knowledge discrepancies about the identity of mutually-visible objects. Our results indicated that in this scenario, children showed notable difficulty in suppressing privileged knowledge. Specifically, as the target word was heard, children showed a tendency to consider a visually-misleading competitor object even when the speaker's belief state would in principle prevent her from labelling that object with a word sharing sounds with the target name. Although this contrasts with children's demonstrated ability to manage privileged knowledge about the objects that are present in the display, the findings are in line with studies of adults showing differences in perspective-taking depending on the precise type of knowledge mismatch (Keysar et al., 2003; Mozuraitis et al., 2015). For example, using similar kinds of visually misleading objects, Mozuraitis et al. (2015) found that adult listeners also show comparatively greater difficulty in suppressing privileged knowledge about objects' identity. The emerging picture is therefore consistent in showing that the way perspective information influences real-time language comprehension differs depending on the type of perspective reasoning involved.

What is it exactly that makes managing privileged knowledge about object identity so difficult (compared to the more well-studied scenario where the listener knows about objects that the speaker doesn't)? We noted earlier that this process is likely to be more cognitively demanding due to the more complex mental representations involved. That is, in a situation where differences in perspectives involve whether a partner can see object X or Y, children can

simply restrict attention to the set of entities mutually visible in the referential domain. In this case, little is needed beyond basic attentional mechanisms (e.g., excluding certain entities from the candidate set). In contrast, mentalizing others' knowledge about object identity requires children to infer the way in which a speaker categorizes a mutually-visible object. Thus, the problem is not one of "domain narrowing" but instead one of considering and actively maintaining the possible ways in which objects could be construed (and, in turn, named).

Consistent with this explanation, studies have demonstrated that the ability to appreciate others' lack of knowledge about certain aspects of objects or states of affairs that they themselves know is challenging for children (e.g., Apperly & Robinson, 1998; Apperly & Robinson, 2003; Kamawar & Olson, 1999, 2009, 2011; Sprung et al., 2007). For example, preschool children and adults can track a person's false belief about an object's location and will immediately anticipate which location they will go to find an object based on that false belief. In contrast, children at this age will fail to show a sensitivity to another person's false belief regarding the property of an object (e.g., a dog-robot that appeared blue to an adult but red to the child; Low & Watts, 2013).

Other related work has suggested there are two distinct systems of perspective-taking (Butterfill & Apperly, 2013; Low et al., 2016; Low & Watts, 2013). On this account, the ability to track what is or is not seen from another person (level 1 perspective-taking) draws on a largely automatic and fast-acting system. In contrast, the ability to represent how another person sees things (level 2 perspective-taking) draws on a distinct system that works slowly and requires cognitive resources. Although this account can explain the results like those described above and found in our studies, it is not clear that the relevant patterns we found necessarily require a two-systems explanation. For example, these same results may be explained by a single mindreading system where the time course and the nature of processing depend on the cognitive demands of

the task (Carruthers, 2015, 2017). Indeed, when cognitive demands of the tasks increase, listeners make more egocentric errors even in situations involving their partner's knowledge discrepancies about the mere visual availability of objects (Zhao et al., 2018). Thus, a number of questions about the architecture of the perspective-taking system remain unclear.

It is important to acknowledge that although shared knowledge between speakers and listeners plays a central role in conversation, it would not be optimal to fully suppress one's privileged knowledge. Indeed, one goal in most conversations is to exchange information and expand the knowledge mutually known to participants. As Stalnaker (2002) pointed out, mutual knowledge about a certain piece of information is not always present but instead it is established by producing an utterance. In addition, some discourse contexts draw on the use of listeners' privileged knowledge to achieve a communicative goal (Brown-Schmidt & Fraundorf, 2015). Therefore, it seems plausible that privileged knowledge is neither completely ignored nor prioritized but is maintained in different ways and to different extents during language comprehension, depending on the type of knowledge discrepancy, the situational context, and other factors (Heller et al., 2016). For example, Hanna et al. (2003) found that adult listeners effectively used a speaker's false belief about the properties of objects to guide their interpretation, but at the same time were influenced by their own perceptual knowledge. Finally, our findings have methodological implications for research on children's communicative perspective-taking abilities. The results of studies involving differences in communicative partners' visual perspective have often been used as a proxy measure to capture the development of perspective-taking skills (e.g., Fan et al., 2015; Nilsen & Fecica, 2011). Although this type of setting has clearly informed our understanding of the use of perspective information during real-time language processing, the present results highlight the risk of relying on a single task to

benchmark performance (Wellman, 2012). We suggest that, by expanding the range of scenarios used to study the management of knowledge discrepancies, we will arrive at a more comprehensive and solid understanding of perspective-taking in childhood.

3.6. References

- Apperly, I., & Robinson, E. (1998). Children's mental representation of referential relations. *Cognition*, 67(3), 287-309.
- Apperly, I. A., & Robinson, E. J. (2003). When can children handle referential opacity? Evidence for systematic variation in 5-and 6-year-old children's reasoning about beliefs and belief reports. *Journal of Experimental Child Psychology*, 85(4), 297-311.
- Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59(4), 457-474.
- Barr, D. J., & Keysar, B. (2002). Anchoring comprehension in linguistic precedents. *Journal of Memory and Language*, 46(2), 391-418.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
<https://doi.org/10.18637/jss.v067.i01>
- Berman, J. M., Chambers, C. G., & Graham, S. A. (2010). Preschoolers' appreciation of speaker vocal affect as a cue to referential intent. *Journal of Experimental Child Psychology*, 107(2), 87-99.
- Berman, J. M., Graham, S. A., Callaway, D., & Chambers, C. G. (2013). Preschoolers use emotion in speech to learn new words. *Child Development*, 84(5), 1791-1805.
- Berman, J. M., Graham, S. A., & Chambers, C. G. (2013). Contextual influences on children's use of vocal affect cues during referential interpretation. *Quarterly Journal of Experimental Psychology*, 66(4), 705-726.

- Brown-Schmidt, S. (2009, Oct). The role of executive function in perspective taking during online language comprehension. *Psychon Bull Rev*, *16*(5), 893-900.
<https://doi.org/10.3758/PBR.16.5.893>
- Brown-Schmidt, S., & Fraundorf, S. H. (2015). Interpretation of informational questions modulated by joint knowledge and intonational contours. *Journal of Memory and Language*, *84*, 49-74.
- Brown-Schmidt, S., Gunlogson, C., & Tanenhaus, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition*, *107*(3), 1122-1134.
- Brown-Schmidt, S., & Heller, D. (2014). What language processing can tell us about perspective taking: A reply to Bezuidenhout (2013). *Journal of Pragmatics*, *60*, 279-284.
- Brown-Schmidt, S., & Heller, D. (2018). Perspective-taking during conversation. *Oxford handbook of psycholinguistics*, 551-574.
- Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, *28*(5), 606-637.
- Carruthers, P. (2015). Perceiving mental states. *Consciousness and cognition*, *36*, 498-507.
- Carruthers, P. (2017). Mindreading in adults: Evaluating two-systems views. *Synthese*, *194*(3), 673-688.
- Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground at the understanding of demonstrative reference. *Journal of verbal learning and verbal behavior*, *22*(2), 245-258.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1-39.

- Davies, C., Lingwood, J., Ivanova, B., & Arunachalam, S. (2021). Three-year-olds' comprehension of contrastive and descriptive adjectives: Evidence for contrastive inference. *Cognition*, *212*, 104707. <https://doi.org/10.1016/j.cognition.2021.104707>
- Degen, J., & Tanenhaus, M. K. (2019). Constraint-based pragmatic processing. *The Oxford handbook of experimental semantics and pragmatics*, 21-38.
- Dink, J., & Ferguson, B. (2018). *eyetrackingR*. In (Version 0.1.8) <http://www.eyetracking-R.com>
- 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*(6), 760-768.
- Fan, S. P., Liberman, Z., Keysar, B., & Kinzler, K. D. (2015). The exposure advantage: Early exposure to a multilingual environment promotes effective communication. *Psychological Science*, *26*(7), 1090-1097.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007, May). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*, *39*(2), 175-191. <https://doi.org/10.3758/bf03193146>
- Ferguson, H. J., & Breheny, R. (2012). Listeners' eyes reveal spontaneous sensitivity to others' perspectives. *Journal of Experimental Social Psychology*, *48*(1), 257-263.
- Flavell, J. H. (1993). The development of children's understanding of false belief and the appearance-reality distinction. *International Journal of Psychology*, *28*(5), 595-604.
- Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology*, *15*(1), 95-120.

- Fussell, S. R., & Krauss, R. M. (1989). The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of Experimental Social Psychology, 25*(3), 203-219.
- Gibbs Jr, R. W., Mueller, R. A., & Cox, R. W. (1988). Common ground in asking and understanding questions. *Language and speech, 31*(4), 321-335.
- Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development, 26*-37.
- Grice, H. P. (1975). Logic and conversation. In *Speech acts* (pp. 41-58). Brill.
- Grodner, D., & Sedivy, J. C. (2011). 10 The Effect of Speaker-Specific Information on Pragmatic Inferences. In *The processing and acquisition of reference* (Vol. 2327, pp. 239-272). MIT Press. <https://doi.org/10.7551/mitpress/9780262015127.003.0010>
- Hanna, J. E., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. *Cognitive science, 28*(1), 105-115.
- Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language, 49*(1), 43-61. [https://doi.org/10.1016/S0749-596X\(03\)00022-6](https://doi.org/10.1016/S0749-596X(03)00022-6)
- Hansen, M. B., & Markman, E. M. (2005). Appearance questions can be misleading: a discourse-based account of the appearance–reality problem. *Cognitive Psychology, 50*(3), 233-263.
- Heller, D., Grodner, D., & Tanenhaus, M. K. (2008). The role of perspective in identifying domains of reference. *Cognition, 108*(3), 831-836.
- Heller, D., Parisien, C., & Stevenson, S. (2016). Perspective-taking behavior as the probabilistic weighing of multiple domains. *Cognition, 149*, 104-120.

- Huang, Y. T., & Snedeker, J. (2013). The use of lexical and referential cues in children's online interpretation of adjectives. *Developmental Psychology, 49*(6), 1090.
<https://doi.org/10.1037/a0029477>
- Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L. R., & Trueswell, J. C. (2000). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of psycholinguistic research, 29*(6), 597-626.
- Kamawar, D., & Olson, D. R. (1999). Children's representational theory of language: The problem of opaque contexts. *Cognitive Development, 14*(4), 531-548.
- Kamawar, D., & Olson, D. R. (2009). Children's understanding of referentially opaque contexts: The role of metarepresentational and metalinguistic ability. *Journal of Cognition and Development, 10*(4), 285-305.
- Kamawar, D., & Olson, D. R. (2011). Thinking about representations: The case of opaque contexts. *Journal of Experimental Child Psychology, 108*(4), 734-746.
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science, 11*(1), 32-38.
- Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition, 89*(1), 25-41.
- Khu, M., Chambers, C., & Graham, S. A. (2018). When You're Happy and I Know It: Four-Year-Olds' Emotional Perspective Taking During Online Language Comprehension. *Child Development, 89*(6), 2264-2281.
- Khu, M., Chambers, C. G., & Graham, S. A. (2020). Preschoolers flexibly shift between speakers' perspectives during real-time language comprehension. *Child Development, 91*(3), e619-e634.

- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(1), 1-26.
<https://doi.org/10.18637/jss.v082.i13>
- Low, J., Apperly, I. A., Butterfill, S. A., & Rakoczy, H. (2016). Cognitive architecture of belief reasoning in children and adults: A primer on the two-systems account. *Child Development Perspectives*, *10*(3), 184-189.
- Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological science*, *24*(3), 305-311.
- Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & psychophysics*, *53*(4), 372-380.
<https://doi.org/10.3758/BF03206780>
- 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.
- Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton, Florida : CRC Press.
- Moll, H., & Tomasello, M. (2012). Three-year-olds understand appearance and reality—just not about the same object at the same time. *Developmental Psychology*, *48*(4), 1124.
- Mozuraitis, M., Chambers, C. G., & Daneman, M. (2015). Privileged versus shared knowledge about object identity in real-time referential processing. *Cognition*, *142*, 148-165.
- Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, *13*(4), 329-336.

- Nilsen, E. S., & Fecica, A. M. (2011). A model of communicative perspective-taking for typical and atypical populations of children. *Developmental Review, 31*(1), 55-78.
- Nilsen, E. S., & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology, 58*(2), 220-249.
- Nilsen, E. S., & Graham, S. A. (2012). The development of preschoolers' appreciation of communicative ambiguity. *Child Development, 83*(4), 1400-1415.
- R Core Team. (2020). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ryskin, R., Kurumada, C., & Brown-Schmidt, S. (2019, Aug). Information Integration in Modulation of Pragmatic Inferences During Online Language Comprehension. *Cogn Sci, 43*(8), e12769. <https://doi.org/10.1111/cogs.12769>
- Ryskin, R. A., Benjamin, A. S., Tullis, J., & Brown-Schmidt, S. (2015). Perspective-taking in comprehension, production, and memory: An individual differences approach. *Journal of Experimental Psychology: General, 144*(5), 898.
- San Juan, V., Khu, M., & Graham, S. A. (2015). A new perspective on children's communicative perspective taking: When and how do children use perspective inferences to inform their comprehension of spoken language? *Child Development Perspectives, 9*(4), 245-249.
- Sapp, F., Lee, K., & Muir, D. (2000). Three-year-olds' difficulty with the appearance–reality distinction: Is it real or is it apparent? *Developmental Psychology, 36*(5), 547.
- Saryazdi, R., & Chambers, C. G. (2021). Real-time communicative perspective taking in younger and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(3).

- Sprung, M., Perner, J., & Mitchell, P. (2007). Opacity and discourse referents: Object identity and object properties. *Mind & Language*, 22(3), 215-245.
- Stalnaker, R. (2002). Common ground. *Linguistics and philosophy*, 25(5/6), 701-721.
- Syrett, K., LaTourrette, A., Ferguson, B., & Waxman, S. R. (2019). Crying helps, but being sad doesn't: Infants constrain nominal reference online using known verbs, but not known adjectives. *Cognition*, 193, 104033. <https://doi.org/10.1016/j.cognition.2019.104033>
- Trueswell, J. C. (2008). Using eye movements as a developmental measure within psycholinguistics. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 73-96). John Benjamins.
- Wellman, H. M. (2012). Theory of mind: Better methods, clearer findings, more development. *European Journal of Developmental Psychology*, 9(3), 313-330.
- Wellman, H. M. (2018). Theory of mind: The state of the art. *European Journal of Developmental Psychology*, 15(6), 728-755.
- Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language*, 31(2), 183-194.
- Zhao, L., Wang, J. J., & Apperly, I. A. (2018). The cognitive demands of remembering a speaker's perspective and managing common ground size modulate 8-and 10-year-olds' perspective-taking abilities. *Journal of Experimental Child Psychology*, 174, 130-149.

CHAPTER FOUR: Conclusions

4.1. Goals and summary of findings

In this dissertation, I investigated children's generation and management of pragmatic inferences during online language comprehension. Specifically, I examined children's online pragmatic inferencing in two situations: i) when inferences were assumed to be comparatively context-independent (Chapter 2) and ii) when inferences relied on more specific aspects of a communicative context (Chapter 3). In both scenarios, I examined whether preschoolers attend to the characteristics of the speaker (the speaker's conventional/unconventional use of language or their knowledge state) when making inferences about the speaker's intended meaning. More specifically, in Chapter 2, I examined whether children adjust their "default" expectations about a speaker's intended referent based on whether the speaker followed or violated norms of communicative behaviour. In Chapter 3, I examined whether the type of knowledge that children share or do not share with a speaker influences how effectively children manage situation-specific information to guide inferences about a speaker's intended meaning. To gain insight into the time course of language processing, children's eye movements were recorded as they heard the spoken instructions. Taken together, the results yield several insights into children's real-time pragmatic inferences in different communicative situations.

In Experiment 1 (Chapter 2), 4- and 5-year-old children were introduced to either a conventional speaker, who used language in expected ways or an unconventional speaker who used language in unexpected ways. That is, the unconventional speaker used verbose and infelicitous descriptions to refer to an object. On critical trials, children were instructed to look at an object (e.g., "Look at the big duck.") when presented with displays containing a target, a competitor, a distractor, and either a contrasting object in the same category as the target object (on Contrast trials) or another distractor (on No contrast trials). When interacting with the

conventional speaker, both 4- and 5-year-old children rapidly generated contrastive inferences during incremental processing. That is, when a contrast object existed on the display (e.g., another duck), children directed their attention towards the target object as they heard the prenominal size adjective (e.g., “big”, in “the big duck”). This reflects the conventional assumption that rational speakers use size adjectives only when needed to differentiate the target from a contrasting alternative, and would otherwise not use a modifier (e.g., simply “the duck”). This assumption also explains the pattern found when the contrast object was replaced with an unrelated object, where children instead showed a tendency to look at the competitor object during the unfolding adjective (i.e., the other big object). Conversely, when interacting with the unconventional speaker, children did not reveal this pattern but showed a delayed preference to look at the target object over the competitor during contrast trials. In other words, contrastive inferencing was attenuated.

In Chapter 3, I examined children’s online pragmatic inferences in a communicative context that involved knowledge discrepancies between children and a speaker. Five-year-old children completed a task with a speaker in a situation where the knowledge mismatch between the children and the speaker was either present or absent. In Experiment 1, children were asked to follow a speaker’s instruction (e.g., “Look at the candy.”) in either a situation where the speaker and child listener shared the knowledge about the actual identity of a mutually visible but visually-deceptive competitor object in a display (e.g., a candle that looks like an apple) or a situation where only the child knows the actual identity of that object. Of interest was the degree to which competitor objects were considered as the target word was heard. In Experiment 2, children’s performance in a scenario like the one in Experiment 1 was compared to a case where the speaker-hearer knowledge mismatch involved objects that were hidden from the speaker. In

this case, only the hearer was aware of certain objects' existence. Results indicated that children showed an egocentric tendency to consider their privileged knowledge about the identities of visually misleading objects and struggled to take the speaker's knowledge state into account. However, when the knowledge mismatch involved private knowledge about the existence of certain objects, children readily took a speaker's knowledge into account to guide their interpretation. This pattern of results suggests children's ability to infer referential intent from the situation-specific "common ground" information is not uniform but depends on the kind of knowledge being tracked.

4.2. General discussion

First, the studies in this dissertation contribute to a growing body of evidence suggesting that children rapidly make inferences about a speaker's intended referent during online language comprehension. The present findings indicate that preschool children use a variety of pragmatic cues in the online interpretative process and offer insight into two specific types of pragmatic computations: 1) those derived from conventional patterns of language use (and are understood to be independent of context); and 2) those reliant on specific features of the communicative context. The specific evidence comes from the finding that 1) children understand that a speaker's use of a certain expression (i.e., a pronominal size adjective) carries pragmatic meaning and use this knowledge to guide their interpretation (Chapter 1, Experiment 1 with the conventional speaker); and 2) children understand that a speaker would not know about the existence of an object hidden from their view and apply this knowledge during referential processing in real-time (Chapter 2, Experiment 2). These outcomes provide additional evidence against the classic proposal that the effort associated with inferencing means children will be

unsuccessful at making these computations in real-time (Epley et al., 2004; Fan et al., 2015; Keysar et al., 1998).

Furthermore, the results of Experiment 1 advance our understanding in important ways by demonstrating children's *flexible* pragmatic inferencing, even in cases where these inferences are thought to rest on conventional patterns of language use. That is, preschool children do not simply react to the speaker's use of certain words but instead consider whether the speaker abides by the norms of communication and make adjustments accordingly. Thus, the communicative norms involved in contrastive inference should not be understood as strict "rules", but instead as behaviours that children attempt to rationalize against the situational context. It is important to note, however, that although children's online contrastive inferencing was delayed in the unconventional speaker condition, they still derived contrastive inferences when hearing referring expressions produced by an unconventional speaker. The fact that the inferences were not fully extinguished might be argued to support the idea that some pragmatic inferences are available without fully engaging in pragmatic reasoning about a speaker's intention or beliefs. Such a view aligns with Andrés-Roqueta and Katsos (2017)'s use of the term linguistic-pragmatic to refer to some cases where listeners can compute pragmatic inferences once their structural language ability and knowledge with pragmatic norms are in place. For example, adolescents with autism spectrum disorder computed scalar implicatures (where the use of a term *some* in a sentence like *Some ducks are white* implies that the stronger alternative *all* is not true) without engaging in the reasoning about a speaker's epistemic state (Hochstein, Bale, & Barner, 2018). However, it is also possible to argue that complete suppression of contrastive reasoning on the basis of the speaker's behaviour is unlikely. Although the speaker clearly departed from communicative conventions by producing long-winded descriptions, this

behaviour does not indicate that the speaker is incapable of producing contrastive descriptions that are appropriate in a given context. The important finding is therefore that children reacted to the unconventional behaviour by changing the nature of their contrastive inferencing and not by suppressing it entirely.

The findings from Chapter 3 offer new insights into how the children's use of a speaker's knowledge in online comprehension varies depending on the type of perspective representation children need to hold. I found that children effectively managed privileged knowledge versus shared knowledge about the physical co-presence of an object while processing a referring expression. In contrast, children's privileged knowledge about the actual *identities* of (visually deceptive) objects was more difficult to manage. Children's referential interpretation was affected by this knowledge even when they were aware of the speaker's lack of knowledge about the deceptive object's identity. This pattern of findings is consistent with previous research with adults (Mozuraitis et al., 2015). Taken together, the adult and child data provide evidence that certain kinds of privileged knowledge strongly impact the efficiency of interpretative processing.

There are two possible explanations for why listeners consider certain types of privileged knowledge when processing referring expressions. First, it might be because different mechanisms or systems are involved with different types of mentalizing (Apperly & Butterfill, 2009; Low & Watts, 2013), or that more demands are placed on certain kinds of mentalizing (Carruthers, 2016, 2017). A second possible explanation, suggested by Mozuraitis and colleagues (2015), proposed that privileged knowledge about the actual identities of objects is closely related to psychological essentialism, which is an intuitive belief that certain categories share some deeper and nonobvious reality (Gelman, 2004). For example, children judge the identity and the function of an object based on its essence, not appearance (Gelman & Wellman,

1991). Therefore, the superficially egocentric tendency to consider one's privileged knowledge of object identity might reflect a bias towards conceptualizing entities in terms of their inherent (rather than superficial) nature.

Together, the findings from my dissertation highlight the importance of considering diverse aspects of pragmatic inferences. Some kinds of pragmatic cues might be difficult to integrate with linguistic information during online interpretation. Understanding a broader range of pragmatic inferences and the development of children's pragmatic competence may also allow us to have better understanding of which cognitive skills need to be acquired. Indeed, there is evidence that distinct kinds of perspective-taking rely on different cognitive skills (Ryskin et al., 2015). Therefore, more work is needed regarding children's pragmatic competence in different situations to understand how their behaviour interfaces with various cognitive capacities.

Lastly, this dissertation adds to the existing literature on children's pragmatic competence by highlighting the advantages of using a remote conversational paradigm. The experiments in this dissertation all involved a methodology in which children were introduced to a person in another room with whom they were going to interact. The experimenter pretended to communicate with the speaker, which led children to believe that they were interacting with the person who they had met in real-time. Interactivity has been suggested to be an important factor in work on real-time perspective-taking in adults (e.g., Brown-Schmidt et al., 2008) and in the developmental literature (e.g., Syrett et al., 2019). It is possible that interactivity might motivate children to engage in a task and thus increase their sensitivity to pragmatic cues. The design also allowed us to avoid the problem where interactive settings can bring about a lack of control in the experimental materials. For example, there can be variability in critical sentences uttered by

different confederate speakers. To avoid these issues, we used pre-recorded materials in the task so that all children heard the same spoken sentences.

The experiments in this dissertation raise some questions to be addressed with future research. First, I found that children still derived contrastive inferences even when they communicated with an unconventional speaker in Chapter 1. This finding raises questions around how this type of pragmatic inferencing becomes automatized. Only 6% of adjectives presented to 3- and 4-year-old children in the UK are contrastive (Davies, Lingwood, & Arunachalam, 2020). It might be that situations where prenominal size adjectives are used contrastively are salient enough for children to learn the pragmatic function of size adjectives with just a few exposures. Future research could examine the contexts in which size adjectives are used in child-directed speech and how these contexts may differ from other prenominal adjectives (e.g., colour).

A second avenue for future research could focus on the type and the amount of evidence required for children to judge a speaker as unconventional. The unconventional speaker in Experiment 1 flouted the maxims of quantity (the amount of information provided in a description) as well as manner (e.g., referring to a banana in terms of an animal that often eats it). Furthermore, children received both top-down (i.e., heard that the speaker says things in a weird way) and bottom-up (i.e., verbose and weird referential expressions) information about speaker conventionality. As such, it is uncertain whether children will show adaptive behaviour when, for example, a speaker fails to provide the appropriate amount of information and/or when they receive only top-down information about a speaker's unconventional behaviour. Future research might focus more selectively on violations of particular maxims, such as examining violations of the maxim of quantity in the context of bottom-up evidence alone.

A third avenue for future research could focus on extent to which privileged knowledge about object identity influences language processing. In Chapter 3, I found that children considered privileged knowledge about the actual identities of deceptive objects during the unfolding noun. However, they eventually landed on and pointed to the target object because the instructions were quickly disambiguated as the final part of the target noun was heard. Therefore, it is not clear whether children fail to appreciate a speaker's perspective, or whether their integration of this perspective is just delayed during referential processing. This question can be addressed by asking children to infer a speaker's intended referent after listening to truncated referential instructions. For example, children hear an audio instruction which ends before the disambiguation point (e.g., "Look at the ca...") when the visual display involves a candy, a candle that looks like an apple, and two other objects. Will children be able to determine the speaker's intended object with the partial information about the noun if prompted to guess what the speaker is referring to? If children consider a speaker's ignorance about object identity, they would choose the candy as the upcoming noun. However, it is possible that children might have global difficulties in inhibiting their knowledge about the deceptive object and thus make an egocentric error of choosing the candle that looks like an apple even when given substantial decision-making time. It would be interesting to examine how much children's language processing is interrupted by the "curse of knowledge" and whether they can inhibit privileged knowledge when there is clear referential ambiguity.

Overall, this dissertation adds to a growing literature that demonstrates children's pragmatic inferencing in real-time language comprehension. More specifically, the findings of this dissertation demonstrate that preschool children consider a speaker's conventionality and knowledge about physical presence of objects to infer the intended referent during online

language processing. Moreover, my research suggests that the way children manage privileged versus shared knowledge differs depending on the type of knowledge applied in the context. By investigating children's pragmatic inferences in various contexts, future research can provide a more comprehensive account of children's pragmatic abilities.

References

- Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247-264.
- Andrés-Roqueta, C., & Katsos, N. (2017). The contribution of grammar, vocabulary and theory of mind in pragmatic language competence in children with autistic spectrum disorders. *Frontiers in psychology*, 8, 996.
- Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological review*, 116(4), 953.
- Barr, D. J. (2008). Analyzing ‘visual world’ eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59(4), 457-474.
<https://doi.org/10.1016/j.jml.2007.09.002>
- Barr, D. J., & Keysar, B. (2002). Anchoring comprehension in linguistic precedents. *Journal of Memory and Language*, 46(2), 391-418.
- Brown-Schmidt, S. (2009, Oct). The role of executive function in perspective taking during online language comprehension. *Psychon Bull Rev*, 16(5), 893-900.
<https://doi.org/10.3758/PBR.16.5.893>
- Brown-Schmidt, S., Gunlogson, C., & Tanenhaus, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition*, 107(3), 1122-1134.
- Carruthers, P. (2015). Perceiving mental states. *Consciousness and cognition*, 36, 498-507.
- Carruthers, P. (2017). Mindreading in adults: Evaluating two-systems views. *Synthese*, 194(3), 673-688.

- Carston, R. (1998). Informativeness, relevance and scalar implicature. *Pragmatics And Beyond New Series*, 179-238.
- Clark, E. V. (2015). 15 Common Ground. *The handbook of language emergence*, 328.
- Clark, H. H. (1992). *Arenas of language use*. University of Chicago Press.
- Clark, H. H., & Marshall, C. R. (1981). Definite knowledge and mutual knowledge.
- Davies, C., Lingwood, J., & Arunachalam, S. (2020). Adjective forms and functions in British English child-directed speech. *Journal of Child Language*, 47(1), 159-185.
- Davies, C., Lingwood, J., Ivanova, B., & Arunachalam, S. (2021). Three-year-olds' comprehension of contrastive and descriptive adjectives: Evidence for contrastive inference. *Cognition*, 212, 104707. <https://doi.org/10.1016/j.cognition.2021.104707>
- Degen, J., & Tanenhaus, M. K. (2019). Constraint-based pragmatic processing. *The Oxford handbook of experimental semantics and pragmatics*, 21-38.
- 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(6), 760-768.
- Fan, S. P., Liberman, Z., Keysar, B., & Kinzler, K. D. (2015). The exposure advantage: Early exposure to a multilingual environment promotes effective communication. *Psychological Science*, 26(7), 1090-1097.
- Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, 72(4), 1003-1015.
- Fernald, A., Thorpe, K., & Marchman, V. A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective–noun phrases. *Cognitive Psychology*, 60(3), 190-217.

- Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology, 17*(1), 99.
- Frazier, L. (1987). Sentence processing: A tutorial review.
- Gelman, S. A. (2004). Psychological essentialism in children. *Trends in cognitive sciences, 8*(9), 404-409.
- Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition, 38*(3), 213-244.
- Grice, H. P. (1975). Logic and conversation. In *Speech acts* (pp. 41-58). Brill.
- Grodner, D., & Sedivy, J. C. (2011). 10 The Effect of Speaker-Specific Information on Pragmatic Inferences. In *The processing and acquisition of reference* (Vol. 2327, pp. 239-272). MIT Press. <https://doi.org/10.7551/mitpress/9780262015127.003.0010>
- Hanna, J. E., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. *Cognitive science, 28*(1), 105-115.
- Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language, 49*(1), 43-61. [https://doi.org/10.1016/S0749-596X\(03\)00022-6](https://doi.org/10.1016/S0749-596X(03)00022-6)
- Heller, D., Grodner, D., & Tanenhaus, M. K. (2008). The role of perspective in identifying domains of reference. *Cognition, 108*(3), 831-836.
- Hochstein, L., Bale, A., & Barner, D. (2018). Scalar implicature in absence of epistemic reasoning? The case of autism spectrum disorder. *Language Learning and Development, 14*(3), 224-240.

- Huang, Y. T., & Snedeker, J. (2013). The use of lexical and referential cues in children's online interpretation of adjectives. *Developmental Psychology, 49*(6), 1090.
<https://doi.org/10.1037/a0029477>
- Huetting, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta psychologica, 137*(2), 151-171.
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science, 11*(1), 32-38.
- Keysar, B., Barr, D. J., Balin, J. A., & Paek, T. S. (1998). Definite reference and mutual knowledge: Process models of common ground in comprehension. *Journal of Memory and Language, 39*(1), 1-20.
- Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition, 89*(1), 25-41.
- Khu, M., Chambers, C. G., & Graham, S. A. (2020). Preschoolers flexibly shift between speakers' perspectives during real-time language comprehension. *Child Development, 91*(3), e619-e634.
- Kukona, A., Fang, S.-Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. *Cognition, 119*(1), 23-42.
- Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. MIT press.
- Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science, 24*(3), 305-311.

- McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*(3), 283-312.
- Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, *13*(4), 329-336.
- Nilsen, E. S., & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology*, *58*(2), 220-249.
- Ryskin, R., Kurumada, C., & Brown-Schmidt, S. (2019, Aug). Information Integration in Modulation of Pragmatic Inferences During Online Language Comprehension. *Cogn Sci*, *43*(8), e12769. <https://doi.org/10.1111/cogs.12769>
- Ryskin, R. A., Benjamin, A. S., Tullis, J., & Brown-Schmidt, S. (2015). Perspective-taking in comprehension, production, and memory: An individual differences approach. *Journal of Experimental Psychology: General*, *144*(5), 898.
- San Juan, V., Khu, M., & Graham, S. A. (2015). A new perspective on children's communicative perspective taking: When and how do children use perspective inferences to inform their comprehension of spoken language? *Child Development Perspectives*, *9*(4), 245-249.
- Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of psycholinguistic research*, *32*(1), 3-23.
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*(2), 109-147. [https://doi.org/10.1016/S0010-0277\(99\)00025-6](https://doi.org/10.1016/S0010-0277(99)00025-6)
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition* (Vol. 142).

- Spivey-Knowlton, M., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, *55*(3), 227-267.
- Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, *71*(2), 73-108.
- Syrett, K., LaTourrette, A., Ferguson, B., & Waxman, S. R. (2019). Crying helps, but being sad doesn't: Infants constrain nominal reference online using known verbs, but not known adjectives. *Cognition*, *193*, 104033. <https://doi.org/10.1016/j.cognition.2019.104033>
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632-1634.
- Tribushinina, E., & Mak, W. M. (2016). Three-year-olds can predict a noun based on an attributive adjective: evidence from eye-tracking. *Journal of Child Language*, *43*(2), 425-441.
- Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, *35*(4), 566-585.
- Weber, A., Braun, B., & Crocker, M. W. (2006). Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language and speech*, *49*(3), 367-392.
- Wolter, L., Gorman, K. S., & Tanenhaus, M. K. (2011). Scalar reference, contrast and discourse: Separating effects of linguistic discourse from availability of the referent. *Journal of Memory and Language*, *65*(3), 299-317.