CHILDREN'S AWARENESS OF SYNTACTIC AMBIGUITY

by

Elly Zimmer

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF LINGUISTICS

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2016

.

THE UNIVERSITY OF ARIZONA GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Elly Zimmer, titled Children's Awareness of Syntactic Ambiguity and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

____ Date: 6 June 2016
Cecile McKee

____ Date: 6 June 2016
Gayle DeDe

____ Date: 6 June 2016
Perry Gilmore

____ Date: 6 June 2016
Heidi Harley

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College. I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Date: 6 June, 2016
Dissertation Director: Cecile McKee

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Elly Zimmer

Acknowledgements

First, I give my thanks to the National Science Foundation for funding Studies 2 and 3 of this dissertation (BCS-1451665). I would also like to thank the schools that allowed me to recruit and test participants at their facilities: Desert Spring Children's Center, Second Street Children's School, St. Mark's Preschool and Kindergarten, Castlehill Country Day School, Green Fields Country Day School, Satori School, Sonoran Science Academy – Tucson, and the International School of Tucson. I would also like to thank these members of the Developmental Psycholinguistics Lab at the University of Arizona for their help designing and conducting this research: Mia Vento, Hui-Yu Huang, Matthew Mutterperl, Thomas McDonough, Margarita Beltran, Sean Harley, Rachel Billeci, Jaqueline Dowell, Holly Durr, and Clare Emmert.

I would like to express my deep gratitude to the members of my dissertation committee for their input on this dissertation: Dr. Gayle DeDe, Dr. Heidi Harley, Dr. Perry Gilmore, and the late Dr. Richard Ruíz. I would like to give a very special thank you to my advisor and committee chair, Dr. Cecile McKee, for her considerable input and guidance on the design and conduct of this research as well as the writing of this dissertation.

Finally, I would like to thank the many people in my life who have supported me throughout my education: my family, Ginger Davis, Willy Zimmer, Emily Zimmer and Will Zimmer; my partner and linguistic sounding board, Noah Nelson; and my dear friends, Abbie Patik, Jenna Milliken, Amber Lebsock, Rikki Peck, Danika Jost, and Gracie Bingham. I could not have done this without them.

TABLE OF CONTENTS

Abstract

This dissertation probes children's metalinguistic awareness of syntactic ambiguity (as in the sentence *The man is poking the monkey with a banana,* where the PP *with a banana* can be understood in two ways, associated with either the monkey or the poking). Several studies suggest that children do not spontaneously detect syntactic ambiguity until the second grade (e.g., Wankoff, 1983; Cairns et al., 2004). However, syntactic ambiguity detection contributes to reading comprehension skills in second and third graders (e.g., Cairns et al., 2004; Yuill, 2009). This research suggests the hypothesis that syntactic ambiguity awareness should contribute to reading development. Specifically, the theoretical model known as the Simple View of Reading posits that the main components of reading are decoding and linguistic comprehension. Syntactic ambiguity detection could contribute to linguistic comprehension because it helps a listener to overcome comprehension difficulties caused by misinterpreting an ambiguous sentence. Thus, it is important to better understand the early development of syntactic ambiguity awareness. If its connection to reading begins younger than second grade, it might be incorporated into early reading curricula and intervention strategies, which are more effective when applied earlier.

This dissertation includes three manuscripts that are or will be submitted for publication. The first manuscript reports on a study that laid the foundation for the following two by testing whether 3- to 5-year-olds access both interpretations of a syntactic ambiguity using a truth value judgment task. The results showed that children do entertain both interpretations, indicating that comprehension should not be an impediment to syntactic ambiguity detection. This study is currently in revisions at *First Language* (Zimmer, 2016a). The second manuscript reports on a study that tested whether 4- to 7-year-olds can detect ambiguous sentences using a task that differs from those used in previous studies. My study used a picture selection task that tested for conscious awareness by having children teach a puppet why multiple pictures could match one sentence. I developed a scoring system for children's explanations that allowed for more gradient measures of early ambiguity awareness than previous research. The results showed that a small proportion of 4- to 7-year-olds are aware of syntactic ambiguity, and many others are beginning to show indications of such awareness (e.g., they select both pictures but their explanations are not yet adult-like). This manuscript is submitted to *The Journal of Psycholinguistic Research* (Zimmer, 2016b).

The third manuscript reports on a study that tested whether 6- to 7-year-olds can learn syntactic ambiguity detection and whether the learning correlates with improvement at reading readiness measures. Participants were divided into two groups: an ambiguity group that did four weeks of games to teach syntactic ambiguity detection, and a control group that did four weeks of math games. I found that children in the ambiguity group improved more at ambiguity detection and at reading readiness tests than those in the control group. This showed that syntactic ambiguity detection is a learnable skill for children as young as 6 and suggests that its connection to reading is in place that young as well. Thus, this skill could be a valuable addition to early reading curricula and intervention strategies. This manuscript will be submitted to *Applied Psycholinguistics* (Zimmer, 2016c).

Chapter 1. Research Questions and Background Literature

1.0 Introduction

This three-article dissertation probes children's metalinguistic awareness of syntactic ambiguity, which likely contributes to reading development (Cairns, Waltzman, & Schlisselberg, 2004; Yuill, 2009). This research approaches syntactic ambiguity detection from several angles to gain a more comprehensive understanding of children's development of this skill. This will help us understand its contribution to reading and explore whether it could be a future component of reading curricula. I collected data in four studies, three of which are described in journal article manuscripts (Appendices A-C).[1] I will refer to these as Studies 1-3. Study 4 is described in Appendix 1 because it did not get beyond piloting. Appendix 2 concerns relationships with research sites like the ones in this dissertation. The best way to read this dissertation is in this order: Chapter 1, Appendices A-C, Chapters 2-3, Appendices 1-2.

Before discussing the research questions for each of the four studies, I will introduce the types of syntactic ambiguity used in this research.

1.1 Three types of syntactic ambiguity

Syntactic ambiguity occurs anytime that one sequence of words can map to multiple syntactic configurations and, thus, has multiple meanings. The focus of this dissertation is global ambiguity, which is never disambiguated by the language itself (Traxler, 2012). This is in contrast to temporary ambiguity, which occurs when a sentence is ambiguous until the addition of a new word limits it to one structural representation (Traxler, 2012). There are several types of global syntactic ambiguity. Studies 1-4 each used a combination of these three types: prepositional phrase attachment ambiguity (PP ambiguity hereafter), reduced relative clause attachment ambiguity (RC ambiguity hereafter), and prenominal adjective attachment ambiguity (adjective ambiguity hereafter).

PP ambiguity occurs when a PP at the end of a transitive sentence can attach either to the verb phrase (VP) as an instrument or the local noun phrase (NP) as a modifier, as shown in (1).

(1) The boy saw the girl <u>with binoculars</u>.[2]

---

[1] Because each article must include its own literature review and methods sections, there will be some repetition of content between the articles and three chapters of this dissertation. The chapters will have expanded discussions that cannot be included in a journal article due to length restrictions. Also, the articles in the dissertation have additional appendices that are not included in the submitted manuscripts due to length restrictions.

[2] Examples are numbered continuously across Chapters 1-3 and Appendices 1-2. Because Appendices A-C are article manuscripts, they each have independent numbering systems.

When the PP *with binoculars* attaches to the local NP, the sentence means that the boy saw a girl who has binoculars. When the PP attaches to the VP, the sentence means that the boy used binoculars to see the girl. The structures for (1) are shown in Figures 1 and 2.

Figure 1: NP-attachment

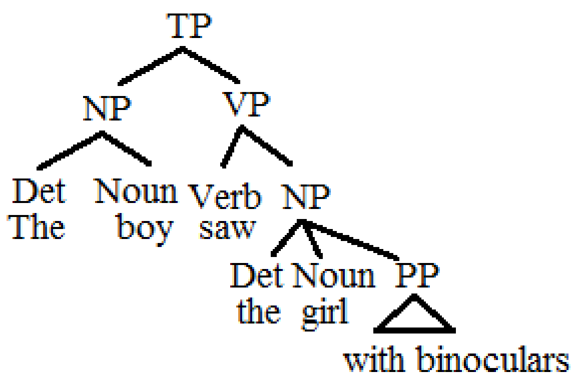
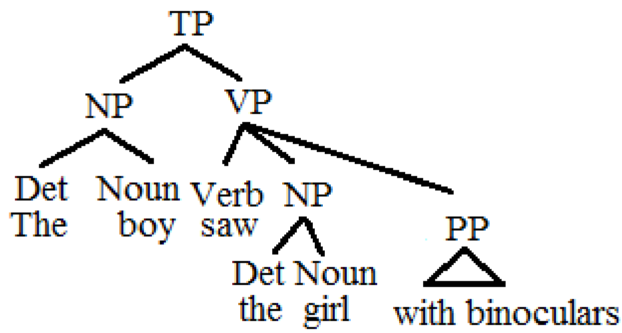


Figure 2: VP-attachment



RC ambiguity has the same attachment sites as PP ambiguity. It occurs when a reduced RC at the end of a sentence can either attach to the VP or to the local NP, as in (2).

(2) The boy saw the girl <u>using binoculars</u>.

When the RC *using binoculars* attaches to the local NP, the sentence means that the boy saw a girl who was using binoculars. When the RC attaches to the VP, the sentence means that the boy used binoculars to see the girl. The key difference between PP and RC ambiguities is that RC ambiguities are bi-clausal, and thus could be considered more complex than PP ambiguities. The structures for (2) are shown in Figures 3 and 4.

Figure 3: NP-attachment

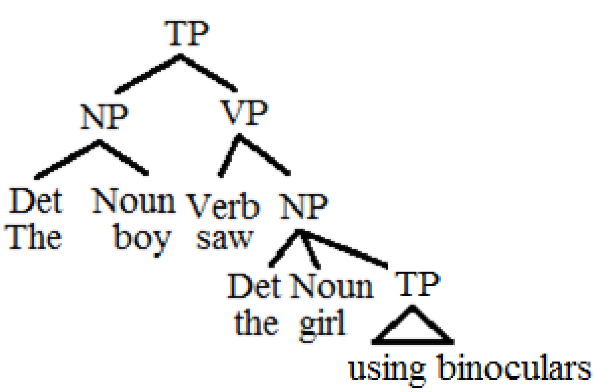
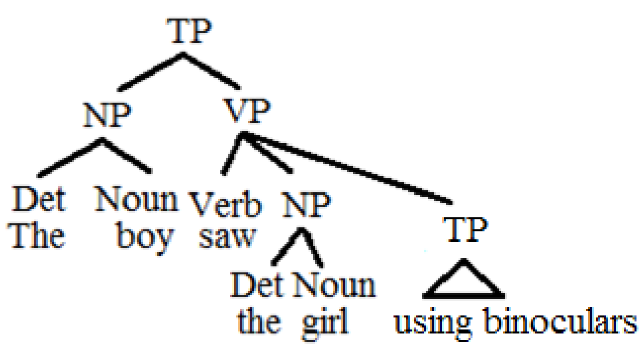


Figure 4: VP-attachment



The third ambiguity type occurs when an adjective precedes a conjoined NP, as in (3).

(3) The teacher found green frogs and fish.

The adjective *green* can modify the conjoined NP *frogs and fish*, so both the frogs and fish must be green, or only the first noun *frogs*, so the fish can be any color. The structures for (3) are shown in Figures 5 and 6.

Figure 5: Only first noun is modified .

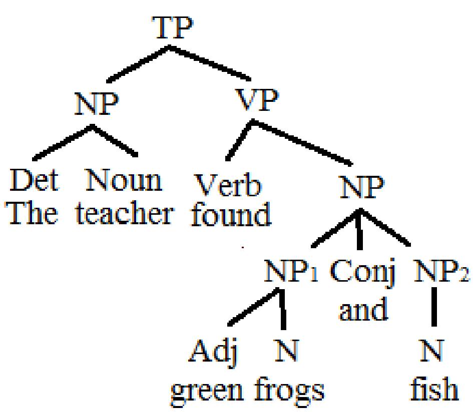
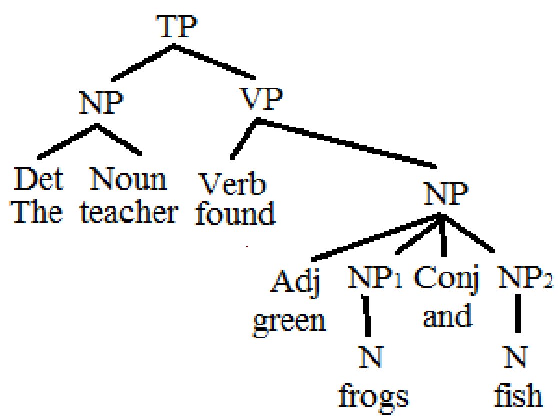


Figure 6: Whole conjoined NP is modified.



2.0 Research Questions

Metalinguistic awareness is generally defined as the ability to consciously think about and manipulate language (Cazden, 1974; Tunmer & Herriman, 1984; Demont & Gombert, 1996; Roth, Speece, Cooper, & De La Paz, 1996; Cairns, 2013). Exactly how metalinguistic awareness relates to the linguistic system is a complex question. A common view is that metalinguistic awareness is a type of metacognition that involves conscious knowledge and control over aspects of the linguistic system (Cazden, 1974; Downing, 1979; Hakes, 1980). Under this view, metalinguistic awareness is not part of the linguistic system per se (Hakes, 1980). There are other views on the relationship between metalinguistic awareness and the linguistic system. For example, Bialystok and Ryan (1985) argue that metalinguistic awareness is not independent from the linguistic system. Rather, it results from the interaction of two separate cognitive factors as they apply to language: cognitive control and analyzed knowledge. When these two cognitive dimensions are mostly inactive, the result is normal conversation. When they are both highly active, the result is metalinguistic awareness. Thus, this framework does not necessarily distinguish metalinguistic abilities from language itself. Rather, they are the output of the same system dependent on the influence of external cognitive factors. Regardless of whether metalinguistic abilities are distinct from language itself, it is clear that metalinguistic awareness cannot occur without language. Thus, it makes sense that the development of linguistic processes would influence the development of metalinguistic abilities. Based on this reasoning, we must first understand the comprehension of syntactic ambiguity before we can understand the development of metalinguistic awareness of syntactic ambiguity. Once we have established this base, we can probe children's conscious awareness of ambiguity taking into account the time course and nature of the development of its comprehension.

Study 1 fills a gap in the literature on children's comprehension of syntactic ambiguity. While several studies have probed children's comprehension of syntactic ambiguity (Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2001; Kidd & Bavin, 2005; Kidd, Stewart, & Serratrice, 2011), their focus has been children's interpretation preferences. To test this, they used versions of tasks, such as act-out and picture selection, that forced children to select only one interpretation. Such research is important as it helps us understand what constraints guide children to choose one interpretation over the other. However, no previous study that I know of has explicitly tested whether individual children can access both interpretations of a single ambiguous sentence. It is important that we know this because accessing both meanings of an ambiguity is a prerequisite to consciously noticing the ambiguity. In other words, if young children cannot comprehend both meanings, they will not detect that a sentence is ambiguous. Thus, Study 1 asked whether children access both interpretations of a syntactic ambiguity. Twenty-six 3- to 5-year-olds participated in a truth value judgment task (Abrams, Chiarello, Cress, Green, & Ellelt, 1978; Crain & McKee, 1985; Gordon & Chafetz, 1986). They judged whether an ambiguous utterance was true after two different stories that illustrated the different interpretations of the sentence. The study included contexts where the same utterances were untrue. Children generally accepted an utterance as true for both contexts supporting its two interpretations but rejected it in a context where it was false. This showed that children can and do access both interpretations. This suggests that, when children fail to notice syntactic ambiguity, it is not because their comprehension system is prohibiting them from accessing both meanings. In this way, Study 1 lays the groundwork for Studies 2 and 3, which probe children's syntactic ambiguity detection.

Before discussing Studies 2 and 3, I will introduce Study 4 because it also focused on a question related to children's comprehension of syntactic ambiguity. Study 4 was designed to fill a gap in the literature on children's parsing of prepositional phrase ambiguities. Previous research has shown that children rely primarily on lexical biases when resolving prepositional phrase ambiguities (Trueswell, Sekerina, Hill, & Logrip, 1999; Snedeker & Trueswell, 2004; Kidd & Bavin, 2005). These biases cause adults and children to prefer one interpretation of an ambiguous construction over another based on the semantic category of the verb in the sentence (Spivey-Knowlton & Sedivy, 1995; Snedeker & Trueswell, 2004; Kidd & Bavin, 2005). However, no study has ever probed the syntactic ambiguity processing of children under 4-years-old. Because lexical biases influence how likely a child is to access both interpretations of a sentence, they will also affect whether or not a child detects the ambiguity (Cairns, 2013). Thus, it is important that we understand the development of these biases in younger children. Study 4 was designed to investigate how lexical biases affect 2- to 4-year-olds' processing of PP ambiguities with an eye-tracking visual world paradigm. The study is described in Appendix 1 because pilot work failed to find an effect of verb bias on looking time in an adult group.

Participants tended to look most at the more complex images, regardless of verb biases. These were usually the images representing a modifier interpretation. This was surprising, given that this effect of verb bias has been found in adults in several studies using other methods, such as act-out tasks paired with eye-tracking and self-paced reading (Spivey-Knowlton & Sedivy, 1995; Trueswell et al., 1999; Snedeker & Trueswell, 2004). The lack of a finding might be due to differences between my task and those used by previous researchers. Specifically, mine was a passive looking task so there was no act-out component. It may be that people look at an array differently when planning out an action than they do when listening passively. For example, it may be that participants in an act-out task look more at possible instruments than they do in a passive looking task. My study did not require any action planning, so participants' looking times may have been more based on visual interest than searching for a plausible instrument. Because the task failed to find any consistent pattern of verb bias among adults, I did not go forward with child subjects. For more information about this pilot study, see Appendix 1.

Unlike Studies 1 and 4, Studies 2 and 3 probed children's metalinguistic awareness directly, asking whether some children develop the ability younger than previously thought. Previous research suggests that syntactic ambiguity detection does not begin to develop until second grade at the earliest (Hirsh-Pasek, Gleitman, & Gleitman., 1978; Shultz & Pilon, 1973; Keil, 1980; Wankoff, 1983; Cairns et al., 2004; Cairns, 2013). However, most previous research used methods that focused on spontaneous detection of syntactic ambiguity. These studies used an interview task where participants were asked explicitly if a set of sentences could have multiple meanings (Shultz & Pilon, 1973; Keil, 1980; Wankoff, 1983; Cairns et al., 2004). For example, Cairns et al. (2004) first gave children a sentence and asked if it meant more than one thing. If the children did not offer a description of the two meanings, experimenters prompted children verbally and then by showing them pictures representing both interpretations. Children's responses received a score on a weighted scale based on when they noticed the ambiguity (i.e., spontaneously, after the verbal prompt, or after the picture prompt). Scores for children under 8-years-old were quite low. It was not until second grade that some children were able to spontaneously detect and describe the syntactic ambiguity. However, I contend that this task and scoring system may have missed indications of syntactic ambiguity awareness in the younger children. Ambiguity detection is not likely to develop at once. It is more likely that it develops over time with aspects of it developing long before a child can spontaneously detect and describe a syntactic ambiguity upon hearing it. Furthermore, there is likely considerable variation between children on the time and rate that they develop this ability, particularly since metalinguistic skills are varied even in adults (Gleitman & Gleitman, 1971; Birdsong, 1989).

Study 2 probed whether some 4- to 7-year-olds show syntactic ambiguity awareness. The 55 children in this study participated in a picture selection task with an explanation component. They chose pictures that matched either an ambiguous or unambiguous sentence. Two pictures matched ambiguous sentences, representing the two meanings of the sentence. One picture matched unambiguous sentences. After picture selection, the children explained their choices to the puppet. Thus, I analyzed two types of data to probe children's awareness: picture choices and explanations. I based my coding system for the explanations on properties of adult explanations of syntactic ambiguity. Rather than focusing on spontaneous detection, this task looks for other indications of ambiguity awareness in the picture selections and explanations, specifically comparing the explanations to an adult target. These finer-grain measures found evidence of early stages of syntactic ambiguity awareness that previous research did not find. The results showed that some 4- to 7-year-olds already have syntactic ambiguity awareness, while others are beginning to develop it. For example, some children chose both pictures for every item but did not give adult-like explanations for their choices. In fact, I found that age did not significantly correlate with how often a child chose both pictures for an ambiguous sentence. However, age did significantly correlate with the quality of children's explanations of syntactic ambiguity when they chose both pictures.

It is important that we understand the development of syntactic ambiguity awareness because some research suggests that it contributes to reading development. A few studies have found correlations between syntactic ambiguity awareness and reading comprehension. Cairns et al. (2004) found that syntactic ambiguity detection in the second grade was correlated with reading scores in the third grade. Yuill (2009) used riddles to train 7- to 9-year-olds to detect structural ambiguities. She found that children who practiced ambiguity detection improved more than a control group and that improvement positively correlated with improvements at reading comprehension. Similarly, Zipke, Ehri, and Cairns (2009) trained third graders to detect lexical and syntactic ambiguity using riddles, games, and passage reading. They found that training did improve participants' ambiguity detection. It also led to improvement on reading comprehension in a paragraph completion task.

This research agenda has two important implications. First, it suggests that syntactic ambiguity awareness does support reading development. Second, Yuill (2009) and Zipke et al. (2009)'s finding suggest that syntactic ambiguity detection is a learnable skill that can be influenced by direct instruction. This is important because, if syntactic ambiguity awareness does support children's reading comprehension and it is teachable, then it could be incorporated into reading curricula and reading intervention measures (Cairns, 2013). However, there remains an important question of whether even younger children can learn to detect syntactic ambiguity. If this skill might help struggling readers, its earlier trainability matters because reading difficulties often appear well before age 8 (Snow, Burns, & Griffin, 1998). Until now, no study has tested whether younger children, who are at earlier yet critical stages of reading development, can be taught syntactic ambiguity detection. Study 3 used a six-week longitudinal design to test whether 6- to 7-year-olds can improve at syntactic ambiguity awareness with practice. I found that children who practiced ambiguity awareness improved significantly more than a control group who practiced math skills. I also found that the ambiguity group's improvement at ambiguity awareness correlated with improvement on reading readiness measures. This suggests that the connection between reading and syntactic ambiguity awareness exists earlier than second grade.

In sum, these studies together contribute to our understanding of children's metalinguistic awareness of syntactic ambiguity. Study 1 laid the groundwork for the other two by showing that

3- to 5-year-olds can access both interpretations of an ambiguity. In Study 2, I found that some children begin to develop syntactic ambiguity awareness younger than previously thought. Finally, in Study 3, I found that 6- to 7-year-olds' syntactic ambiguity awareness can improve with practice and further probed its connection to reading development. The following section provides an overview of the relevant literature for all four studies.

3.0 Background literature

My dissertation research builds on work in three distinct areas of study: parsing, reading, and ambiguity detection. The latter is a small domain, and I will consider it in connection with reading. I begin with comments on the parsing literature because parsing both interpretations of a syntactic ambiguity is the foundation for consciously noticing such an ambiguity.

3.1 Parsing literature

Studies 1-3 did not investigate parsing per se. However, the literature on children's parsing of syntactic ambiguity is very relevant to this research. As mentioned above, metalinguistic abilities cannot exist without language. Thus, features of the linguistic system could affect the development of metalinguistic abilities. Specifically, in the case of syntactic ambiguity detection, if a person does not parse both structures of an ambiguous sentence, it is unlikely that that person would be able to detect the ambiguity (Cairns, 2013). In other words, a hearer must parse both structures of an ambiguity in order to consciously consider both meanings. Thus, it is important to understand children's parsing, particularly for those who the literature claims do not yet have syntactic ambiguity awareness, to address possible reasons why children may fail to detect ambiguity.

Previous research has emphasized differences between adults' and children's processing of syntactic ambiguity. One such difference is that children are less likely to revise their first parse than adults are (Trueswell et al., 1999; Snedeker & Trueswell, 2004; Kidd et al., 2011; Omaki & Lidz, 2015). Trueswell et al. (1999) used eye-tracking during an act-out task to test 4- to 5-year-olds' processing of sentences like the one in (4).

(4) Put the frog on the napkin in the box.

This sentence is temporarily ambiguous until the second PP is uttered. Children's errors in the act-out portion of the task suggested that they did not overcome their initial parses in the same way that adults did. Specifically, children acted out the sentence in non-target ways such as putting the distractor animal on the napkin and hopping the distractor animal to the napkin first and then to the box. This suggests that children parsed the first PP as a destination and could not revise that parse even after hearing the second PP. Kidd et al. (2011) found similar results for globally ambiguous sentences using an act-out task paired with eye-tracking. This matters to ambiguity awareness because if a child fails to revise their initial parse of an ambiguous sentence, he or she is very unlikely to detect the ambiguity. Thus, if 4- to 5-year-olds do not typically revise a parse, this could be part of why they have failed to detect ambiguity in previous research.

Another difference between adults and children's parsing is that children also use fewer informational cues during ambiguity processing than adults (Trueswell et al., 1999; Kidd & Bavin, 2005). Adults draw on several types of informational cues such as referential information, plausibility, and lexical biases (Altmann & Steedman, 1988; Chambers, Tanenhaus, & Magnesun, 2004; Spivey-Knowlton & Sedivy, 1995). However, children apparently rely primarily on one clue: lexical bias. Lexical bias occurs when a person has a preference for one interpretation of a sentence over the other due to the lexical co-occurrence frequencies of the words in the sentence. For example, for PP ambiguities, several studies have shown that verb bias largely drives children's parsing (Trueswell et al., 1999; Snedeker & Trueswell, 2004; Kidd & Bavin, 2005; Kidd et al., 2011). Specifically, the semantic category of action verbs, such as *poke*, more frequently co-occur with VP-attachment. Thus, people tend to parse a PP in a sentence with an action verb as having VP-attachment. The semantic category of perception verbs, such as *listen*, more frequently co-occur with NP-attachment. Thus, people tend to parse a PP in a sentence with a perception verb as having NP-attachment.

Snedeker and Trueswell (2004) used eye-tracking with an act-out task to probe 5-year-olds' sensitivity to verb biases. They found that children's processing was not influenced by referential scene when the verbs were biased towards VP-attachment or NP-attachment. They found only a very weak effect of referential context when verbs were neutrally biased. This suggests children were relying primarily on verb bias during parsing. Similarly, Kidd et al. (2011) paired eye-tracking with an act-out task to probe the interaction of instrument plausibility and verb biases in 4- to 5-year-olds' processing of sentences such as in (5).

(5) Cut the cake with the candle.

The verb *cut* is biased towards VP-attachment, or the instrument interpretation, of a following PP as in (5). This informational cue conflicts with instrument plausibility, as a candle is an implausible instrument for cutting a cake but a very plausible modifier for a cake. Children's responses in the act-out task aligned with verb biases and generally ignored plausibility (i.e., they used the candle to cut the cake in this case). This research supported the hypothesis that 4- to 5-year-olds depend strongly on verb bias. On the other hand, Snedeker and Yuan (2008) did find that 4- to 6-year-olds are somewhat affected by prosody during parsing of PP ambiguities. Together with Snedeker and Trueswell's finding that children show a weak effect of referential context in the absence of a strongly biased verb, this suggests that children can integrate more than one informational cue in their parsing. They may just tend to rely on a smaller set of cues than adults do. Then, as children get older, they begin to incorporate more information types. For example, Kidd and Bavin (2005) showed that 9-year-olds are beginning to attend to noun definiteness during syntactic ambiguity processing. Other cognitive factors also affect children's parsing. Felser, Marinis, and Clahsen (2003) used a self-paced listening task to show that 6- to 7-year-olds' memory span affects their attachment of relative clauses in a globally ambiguous sentence. These various studies paint a picture of the development of ambiguity processing, with children relying predominantly on co-occurrence biases early on and developing the ability to integrate more informational cues in the early school years. Why children rely on lexical biases early on is still an open question, in my view. One possibility comes from Snedeker and Trueswell (2004), who argue that verb bias emerges early because it is more statistically reliable than other cues like referential information.

Children's limited use of informational cues is important to syntactic ambiguity awareness. If pre-school aged children are very biased by the lexical co-occurrence frequencies of the verb, then that may make them less likely to revise a parse and thus to detect an ambiguity. For example, an adult might be likely to notice an ambiguity in a case where her lexical biases push her towards one interpretation but plausibility or the discourse context pushes her towards another. However, with children whose parsing is primarily guided by one cue it is unlikely that extralinguistic informational cues like plausibility would drive them to revise a parse, making it less likely that they would notice an ambiguity.

3.2 Ambiguity detection and reading literature

The literature on children's syntactic ambiguity detection so far suggests that children do not spontaneously detect syntactic ambiguity until 8-years-old at the earliest (e.g., Wankoff, 1983; Cairns et al., 2004). I argue that syntactic ambiguity awareness may begin to develop earlier in at least some children. One reason to anticipate this comes from work on other types of syntactic metalinguistic awareness. For example, it was thought for many years that the ability to judge the well-formedness of sentences did not develop until the first or second grade (Scholl & Ryan, 1980; Tunmer, Pratt, & Herriman, 1984). Yet, research by Dana McDaniel and her colleagues has shown that children as young as 3 can judge sentence grammaticality if they understand that they are discussing language (e.g., McDaniel & Cairns, 1990; 1996; Cairns, Schlisselberg, Waltzman, & McDaniel, 2006). Another type of syntactic awareness is the correction of ungrammatical sentences. Nation and Snowling (2000) showed that children as young as 6 can do this. Such research establishes that young children do have metalinguistic awareness of their syntactic system, suggesting that they may be able to detect syntactic ambiguity.

The second reason to expect that syntactic ambiguity detection may begin for some children at an earlier age relates to the variability of metalinguistic abilities. It is unlikely that children suddenly develop the ability to spontaneously notice an isolated syntactic ambiguity. There are likely aspects of syntactic ambiguity awareness in place well before that type of spontaneous detection occurs. When and how these earlier indications of syntactic ambiguity awareness develop could be highly varied from child to child. Metalinguistic skills tend to be quite variable, even in adults (Gleitman & Gleitman, 1971; Birdsong, 1989). In addition to this natural variability, there are several extralinguistic factors that affect metalinguistic awareness and likely contribute to syntactic ambiguity detection. The number of languages a person speaks is one such factor. Multilinguals outperform monolinguals on many metalinguistic skills (Aronsson, 1981; Bialystok, 1986; 2001b; Davidson, Raschke, & Pervez, 2010; Hermanto, Moreno, & Bialystok, 2012; Bialystok & Barac, 2012). This is in part due to the fact that bilinguals must attend to their languages more in order to switch between them. This is supported by research indicating that bilingual children who translate for their parents have particularly advanced metalinguistic skills (Raschke, 2013) and that bilinguals who are literate in both languages have better metalinguistic awareness than those who are only literate in one language (Thomas, 1988). Both of these facts suggest that bilinguals who spend more time consciously attending to their two languages develop better metalinguistic awareness.

Another factor that affects metalinguistic awareness and relates to bilingualism is executive function, which is the set of cognitive abilities that govern our thought processes. These include interference suppression which allows us to ignore irrelevant information, inhibition of incorrect responses, and attentional control (Traxler, 2012). Some research suggests that bilinguals have advanced metalinguistic skills, in part, because they have enhanced executive function which helps them to isolate and focus on the relevant aspects of their language (Bialystok, 1992; 1999; 2001a; Raschke, 2013). Another extralinguistic factor related to metalinguistic awareness is conservation, which is recognizing that an amount stays the same when it changes container or shape. Research indicates that there are positive correlations between children's proficiency at conservation and their metalinguistic abilities (Hakes, 1980; Wankoff, 1983; Schlisselberg, 1988; van Kleeck, 1994). Thus, it is likely that children will be quite varied in their ability to detect syntactic ambiguity based on both natural variation and these other influencing factors.

In my opinion, we need to understand children's syntactic ambiguity awareness in order to elucidate its likely connection to reading development. For some children, reading difficulties begin as early as pre-school (Snow et al., 1998; Torgesen, 2002; Lonigan & Phillips, 2016). This is an important problem because early reading success predicts later reading success (Juel, 1988; Stanovich, 1986). In fact, most children who are behind on reading in the first few years of school never catch up to their peers without specific reading intervention (Snow et al., 1998; Simmons et al., 2008; Lonigan & Phillips, 2016). There is also a well-known drop-off in reading development around the fourth grade for many students who had previously been showing typical reading skills (Chall, Jacobs, & Baldwin, 1990; Stockard, 2010). This "fourth grade slump" has been attributed to several factors such as the shift from *learning to read* to *reading to learn* (Chall et al., 1990) and vocabulary size (Stanovich, 1986; Graves & Slater, 1987). Reading difficulties, including the "fourth grade slump," are correlated with several socioeconomic factors. Children from many minority groups perform worse on reading assessments than their white counterparts do (Vanneman, Hamilton, Anderson, & Rahman, 2009; Hemphill & Vanneman, 2011). For example, Hispanic students received an average score of 248 on a 2009 national reading assessment, compared to an average score of 271 for their white counterparts (National Center for Education Statistics, 2011). This gap has persisted for several decades.

Similarly, children from families with a low socioeconomic status fall behind their peers in reading as well as in math and science (Snow et al., 1998; National Center for Education Statistics, 1999; Dickinson & Tabors, 2001; Hoff, 2013). Fortunately, intervention strategies can be very effective at closing such achievement gaps if administered early (Duncan & Sojourner, 2013).

It should be noted that there is some disagreement about whether or not we can truly identify reading problems in very young children. Some researchers and educators argue that children are not mentally mature enough to learn to read until around 7-years-old (Moore & Moore, 1975; Elkind, 1987; 2001). They argue that formal reading education in pre-school offers no long-term advantage and may even be a detriment. They cite European nations like Finland, which is currently ranked first in the world in literacy, that do not teach children to read until around 7-years-old (Miller & McKenna, 2016; Andere, 2014). While Finland does have an excellent school system, there is a confound in citing Finland as evidence that children should not learn to read until 7-years-old. Specifically, Finland has an important advantage in terms of reading education in that it has a perfectly transparent alphabet (Ziegler, Bertand, Tóth, Csépe, Reis, et al., 2010). I disagree with the viewpoint that reading skills should not be assessed early as several studies have shown that early reading skills predict reading success in the future (e.g., Wagner & Torgesen, 1987; Snow et al., 1998; Torgesen, 2002; Juel, 1988). For example, in a longitudinal study, Juel (1988) showed that children who struggled with phonological awareness in the first grade were still behind their peers in decoding skills in the fourth grade.

One way to improve intervention strategies for children who need them is to identify the subcomponents of reading. Research has identified several cognitive skills that contribute to reading development such as working memory (Gathercole, Tiffany, Briscoe, & Thorne, 2005; Welsh, Nix, Blair, Bierman, & Nelson, 2010), attentional control (Welsh et al., 2010), and general intelligence (see Stanovich, Cunningham, & Feeman, 1984, for an overview). There are also several metalinguistic skills that we know correlate with reading development. There is some debate on whether metalinguistic skills improve reading or whether reading improves metalinguistic skills (Liberman, Shankweiler, Liberman, Fowler, & Fisher, 1977; Castles & Coltheart, 2004). Some researchers have argued that reading improves metalinguistic awareness (Ehri, 1989; Morais, 1991; Morais, Alegria, & Content, 1987). For example, Morais, Cary, Alegria, & Bertelson (1979) showed that illiterate adults had metalinguistic skills weaker than those of adults who had been illiterate and then learned to read. However, the fact that certain metalinguistic skills in pre-readers correlate with future reading success is strong evidence that metalinguistic abilities contribute independently to reading development (Wagner & Torgesen, 1988). Thus, I conclude that reading and metalinguistic skills both influence each other (Wimmer, Landerl, Linortner, & Hummer, 1991).

One metalinguistic precursor to literacy is phonological awareness (Goswami & Bryant, 1990; Bentin, Deutsch, & Liberman, 1990; Ehri, Nunes, Stahl, & Willows, 2001; National Reading Panel, 2000). Phonological awareness is the ability to consciously identify and manipulate language sounds. It helps the reader decode symbols in the orthography to sounds of the language (Ehri et al., 2001; National Reading Panel, 2000), and is a very important part of early reading development (see Goswami & Bryant, 1990; Castles & Coltheart, 2004 for review papers). Because of this, phonological awareness has become an important component of many reading curricula, most notably phonics-based programs (Baumann, Hoffman, Moon, & Duffy-Hester, 1998). While there is somewhat less research on syntactic awareness, there is evidence that it also contributes to reading development (Tong, Deacon, & Cain, 2014). For example, the ability to identify and correct ungrammatical sentences positively correlates with reading comprehension (Bentin et al., 1990; Demont & Gombert, 1996; Nation & Snowling, 2000).

Syntactic ambiguity awareness is my focus here. It is another type of syntactic awareness, like the ability to make grammaticality judgements. There are both theoretical and empirical reasons to think that syntactic ambiguity awareness also supports reading development. I will begin with theoretical considerations. The Simple View of Reading (SVR) is a well-known model of reading comprehension that proposes two factors that determine reading ability: decoding and linguistic comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990; Joshi & Aaron, 2000; Adlof, Catts, & Little, 2006; Tilstra, McMaster, Van den Broek, Kendeou, & Rapp, 2009). This model defines linguistic comprehension as the ability to derive an interpretation from a sentence <u>and</u> integrate it into a discourse. Proponents of the SVR do not claim that nothing else contributes to reading comprehension. Nor do they claim that decoding or linguistic comprehension are themselves simple processes. They do posit, however, that most of the processes involved in reading comprehension can be subsumed under these two domains. Importantly, these components are dissociable. Children can have poor decoding skills but normal linguistic comprehension, as exemplified by many children with dyslexia (Adlof et al., 2006). Alternately, children can have normal decoding skills but poor linguistic comprehension, as exemplified by children with hyperlexia, which occurs when a person with impaired intellectual abilities demonstrates precocious single-word reading abilities (Gough & Tunmer, 1986; Grigorenko, Klin, & Volkmar, 2003). Thus, these two components are separate, but they are considered equally important in reading development.

Helen Smith Cairns and colleagues suggest that the ability to detect structural ambiguity falls under linguistic comprehension in the SVR (Cairns et al., 2004; Wankoff & Cairns, 2009; Cairns, 2013). Their reasoning goes as follows. If an early reader encounters an ambiguous sentence in a text and parses an interpretation that does not fit the rest of the passage, it will lead to comprehension problems. On the other hand, if the reader notices the ambiguity and switches interpretations, comprehension is less likely to be impaired. Thus, syntactic ambiguity detection should facilitate reading comprehension. This prediction is supported by the previously described studies that found correlations between syntactic ambiguity detection and reading development (Cairns et al., 2004; Yuill, 2009; Zipke et al., 2009). This research suggests that syntactic ambiguity detection could and likely should be incorporated into reading curricula and reading intervention measures (Cairns, 2013). In particular, there are children who struggle with reading comprehension but not with decoding who may benefit from a skill that supports linguistic comprehension directly or uniquely (Stothard & Hulme, 1992; Cain, Oakhill, & Lemmon, 2005;

Grigorenko et al., 2003; Kendeou, Savage, & Van den Broek, 2009). Training on syntactic ambiguity awareness might be a valuable intervention strategy for such children. However, there is much we do not know about the development of this skill and the specific ways that it connects to reading. There is considerable research that needs to take place before we can reach the stage of practical application in curricula and intervention programs.

For example, we need to know more about whether children in the earliest stages of reading can be trained to detect syntactic ambiguity. Yuill (2009) and Zipke et al. (2009) showed that 7- to 9-year-olds can be trained to better detect lexical and syntactic ambiguity. This makes sense, as we already know that 8- to 9-year-olds are beginning to detect syntactic ambiguity on their own (Cairns et al., 2004). However, it is important to examine whether training affects younger children for whom the ability is absent or in earlier stages, especially if we are interested in syntactic ambiguity awareness as a part of future reading intervention strategies. As previously mentioned, children begin to develop reading difficulties as early as preschool. Early reading interventions are more effective than later ones (Vellutino, Scanlon, & Tanzman, 1998). Thus, it matters that we know if syntactic ambiguity awareness can be taught to young children, and if so, if that correlates with improved reading. Two studies provide weak evidence that children under 8-years-old can be taught syntactic ambiguity awareness. Shakibai (2008) and Zipke (2011) showed that children as young as 5 and 6 could be trained to detect <u>lexical</u> ambiguities. Thus, if 5-year-olds can learn to detect one type of linguistic ambiguity, they may be able to learn to detect another.

In sum, this overview of the background literature on children's parsing and metalinguistic awareness of syntactic ambiguity underscores the need for more research in both these domains. Understanding children's syntactic ambiguity awareness is important because it may have practical applications in literacy curricula and interventions for struggling readers. However, before we can develop practical applications, we must better understand how children develop syntactic ambiguity awareness and whether it is a learnable skill for young children. It is also important that we better understand children's comprehension of syntactic ambiguity because it likely affects children's early abilities to detect syntactic ambiguity.

Chapter 2: Further Discussion of Methods

1.0 Introduction

This chapter has two primary goals. The first is to give a comprehensive overview of how the methods of Studies 1-4 compare to and complement each other. The second goal is to give a more in-depth discussion of the research methods employed in this dissertation than is possible in the articles. Because of journal length restrictions, certain details are omitted from the articles that are nevertheless relevant to a dissertation. This includes in-depth consideration of the strengths and weaknesses of each method. I begin with a quick overview of the studies' methods and how they relate to each other. Then, I motivate the methods for each study describing the strengths and weaknesses of each task and data type.

To paint a comprehensive picture of children's metalinguistic knowledge of syntactic ambiguity, we need several different methods. Study 1 and Study 4 were designed to contribute to our understanding of children's ambiguity comprehension in order to better understand the way linguistic processes contribute to metalinguistic awareness. To do this, Study 1 used a truth value judgement task to show that 3- to 5-year-olds access both interpretations of a syntactic ambiguity. Study 4, which is reported in Appendix 1, was a pilot study developed to test how verb biases affect 2- to 4-year-olds' parsing of syntactic ambiguity using a visual-world eye tracking task. Truth value judgement data and eye tracking data matter to the line of research described in this dissertation in that they strengthen our understanding of the comprehension that underlies ambiguity detection. It is also important to note that Studies 1 and 4 focused on 2- to 5-years-olds, children far younger than the age that previous research had concluded that syntactic ambiguity detection begins to develop. I chose preschool children for two reasons. First, there were gaps in the relevant literature regarding these younger ages. More importantly, by focusing on young children's comprehension, we can probe possible reasons why these children have failed to detect syntactic ambiguity in previous research (e.g., Cairns et al., 2004). Because metalinguistic awareness depends on language (Hakes, 1980; Bialystok & Ryan, 1985), it makes sense that certain linguistic competencies must be in place before metacognition of language can develop. Thus, it is important to investigate the linguistic precursors to syntactic ambiguity awareness that develop before awareness itself can develop.

Studies 2 and 3 are different from Studies 1 and 4 because they probed ambiguity detection directly. Thus, they focused on somewhat older children. Because previous research concluded that syntactic ambiguity detection begins at age 8, Study 2 probed syntactic ambiguity detection in 4- to 7-year-olds and found that a small proportion of children this age can detect syntactic ambiguity. To do this, the study used a variation of a picture selection task that collected both comprehension and production data. The comprehension data gave a weak indication of ambiguity awareness when a child chose both pictures representing the two interpretations. The production data provided stronger evidence for conscious knowledge of the ambiguity. The task used in this study is distinct from that in previous research on children's

.

ambiguity detection, which used only production data without the comprehension element (Shultz & Pilon, 1973; Keil, 1980; Wankoff, 1983; Cairns et al., 2004). Further, the production data in these earlier studies seems to have been less specific, having focused on the question of whether a sentence could mean two things rather than analyzing how children explain ambiguity. Having both comprehension and production data in my research gives a more nuanced view of syntactic ambiguity awareness.

Study 3 approached syntactic ambiguity detection from yet another angle, testing whether 6- to 7-year-olds can be taught to better detect syntactic ambiguity and whether that improvement correlates with improvement on measures of reading readiness. To do this, Study 3 used a longitudinal training[1] design that lasted six weeks. The study compared a test group who received ambiguity training, to a control group, who received math skill training. This allowed me to compare a pre-training state to a post-training state. The pre-tests and post-tests consisted of the ambiguity awareness task from Study 2 and reading readiness tests. Thus, I collected comprehension, production, and reading readiness data to measure individual and group differences from before and after the training period. This type of longitudinal data allows for a direct measure of children's learning of syntactic ambiguity awareness. The study found that the participants in the ambiguity group did learn syntactic ambiguity detection and this learning correlated with improvement on reading readiness measures.

In sum, a wide variety of data types was necessary for a nuanced and complete picture of children's development of syntactic ambiguity detection. Using varied methods, I probed children's comprehension and processing of ambiguity to better understand their detection of ambiguity and their ability to learn ambiguity detection. In the following sections, I will give a more detailed description of the strengths and weaknesses of each task and data type used in the three studies reported in the article manuscripts (Appendices A-C). I give some details on how I adjusted my methods to compensate for those weaknesses. One focus is the ways in which the tasks are appropriate for a child population. Besides methodological issues, another very important aspect of working with school-aged children is creating and maintaining relationships with schools where subjects are recruited and research activities are done. For more information on this latter issue, see Appendix 2.

2.0 Study 1 – Truth value judgment task

Study 1 used a truth value judgment task (TVJ hereafter) to test whether 3- to 5-year-olds entertain both interpretations of a syntactic ambiguity. The TVJ task was first used by Abrams et al. (1978) and further developed by Crain and McKee (1985) and Gordon and Chafetz (1986). In

---

[1] In Article 3, I referred to this as *practice* because it is a more common term in education research. Because the focus of this chapter is methods, I will refer to it as *training*, which is the more common methodological term.

a TVJ task, the child hears a story or preamble and then judges whether a set of utterances are true or not based on the story. In other words, children judge the acceptability of an utterance/meaning pair. Thus, this task allows the researcher to test children's interpretation of a particular syntactic construction paired with a particular semantics. It has been used to probe children's comprehension of several structures including passives (Gordon & Chafetz, 1986), quantification (Crain et al., 1996), anaphora (Crain & McKee, 1985), and more. However, it is only appropriate for testing constructions that can be embedded in statements, not questions (Gordon, 1996). This task was appropriate for Study 1's research question of whether children access both meanings of an ambiguous utterance because it allows the researcher to manipulate the meanings of a single utterance via the story contexts. It can also be used to test whether different utterances can be interpreted the same by holding the story context constant and rotating the utterance. In either case, it is important to include filler utterances to mask the purpose of the experiment and avoid priming effects. A manipulation of the task relates to the child's response to the puppet. Children in Study 1 were instructed to tell the puppet "no" if the statement did not match the story and "yes" if the statement did. This is the yes/no variant of the task. The other variant rewards or punishes the puppet by feeding him a cookie if he says the right thing or a rag if he says the wrong thing (Gordon, 1996).

The TVJ task has several strengths. For instance, it is an easy task for children in two ways. First, it does not call on metalinguistic awareness as tasks like grammaticality judgement do (Gordon, 1996). This is important for Study 1 as it is explicitly testing comprehension that occurs in the absence of metalinguistic awareness. The TVJ is also simple in that it has a binary yes/no response, one that children are often called on to make outside of experimental research (e.g., when a parent asks a child whether something is true). This makes the task very simple and constrained for children. Another strength of the task is that it does not feel like a testing situation to children. The puppet's statements are in question, not the children's responses. Thus, children tend to be more comfortable in the task than in some others (Crain, 1998). If researchers make the stories and the puppet's character interesting, the task can be very enjoyable to children. This allows for longer testing sessions than other methods that are less engaging to children. Both its simplicity and the fact that is enjoyable to children, makes the TVJ task suitable for children as young as 2-years-old (Crain & McKee, 1985).

The TVJ also has certain weaknesses. The first relates to people's "yes" bias, which is the tendency to respond "yes" more frequently than "no" in times of uncertainty or pressure, such as an experiment or interview. This tendency is stronger in children than adults, and is exacerbated when the child is unfamiliar with the adult experimenter (Steffensen, 1978; Peterson, Dowden, & Tobin, 1999; Moraguchi, Okanda, & Itakura, 2008). In Study 1, I took several measures to allay this problem. First, I spent time at the schools where the study took place in advance to get to know the child participants. This meant that they were familiar with me before participating in the experiment. Second, the experimental utterances were uttered by a puppet, as children are more comfortable telling a puppet "no" than an adult. A second weakness of the TVJ is the slowness of the task. Because items must be accompanied by story contexts each individual item requires more time than in some other tasks, such as picture selection or act-out task. Because experiments with children must be relatively short due to children's limited attention spans, this means that the TVJ task often allows for fewer total items than other tasks.

The TVJ also requires close attention to several design issues. For example, when developing the contextual stories, it is important to keep them as brief as possible to avoid overloading children's working memory. The stories in Study 1 were extensively piloted to make them as short as possible while still retaining all necessary detail. In addition, the toys used for the stories were positioned to provide visual support to help the child remember the story. For example, in a story where a boy looked at the girl who had binoculars, the toys were left positioned such that the boy was facing the girl and the girl was holding binoculars even after the story was finished. This left the scenario so that the child could check it against the meaning of the sentence, somewhat like a picture. The TVJ task also requires a considerable amount of practice on the part of the experimenters in order to deliver the stories and experimental items in a consistent manner. In fact, this task may require more such practice than almost any other task used to test children's syntactic knowledge. For Study 1, all researchers who tested children practiced the delivery of the stories and experimental items for several hours.

Another design consideration relates to pragmatic factors. Crain et al. (1996) were the first to point out that some research using the TVJ task had included items that were not pragmatically felicitous because they failed to attend to plausible deniability. Plausible deniability stems from the notion that people avoid saying "no" if there is any cause for uncertainty. For example, participants have difficulty judging a statement as false if it introduces a new element into the discourse. This is because participants assume that new discourse elements are introduced for a reason, so introducing such an element in a false statement causes confusion. Thus, in Study 1, for it to be felicitous to a participant to reject one of the puppets' propositions as untrue, all the elements of that proposition should have already been introduced in the discourse. For example, consider the experimental item in (6).

(6) The doctor ate chocolate cakes and cookies.

The story for the false version of this item involved a doctor eating cream-filled cookies and lime-flavored cakes. However, a chocolate bar was also introduced as belonging to one of the other characters in the story. This was so that the statement that participants were judging was not the first mention of chocolate. This was added during piloting, because participants were taken aback when responding to items like the one in (6) when chocolate had not been introduced in the previous discourse. I will now address the strengths and weaknesses of truth value judgement data.

2.1 Truth value judgment data

Truth value judgement data is distinct from conventional comprehension data in important ways. In comprehension tasks, such as the picture selection task or the act-out task, children's responses more directly reflect their interpretation of the sentence. For example, in a picture selection task, when a child chooses a picture, we can infer that the picture represents his interpretation of the sentence. Truth value judgements are a more indirect measure of a child's interpretation because the child is not reporting on his interpretation. Rather, the child is judging whether a sentence is true given a particular context. TVJ data does inform us about participants' interpretation of sentences, though, because children can only judge a sentence as true if they have comprehended it.

TVJ data has an important strength from the researcher's perspective: It is easy to code. Some other comprehension tasks can elicit ambiguous responses from subjects. For example, in the act-out task, children hear an utterance and then act it out with props. Their responses can be difficult to interpret because there are many reasons children could act a sentence out in a non-target way (e.g., they like certain props better than others). However, with the TVJ task, children are supplied the supporting context and the utterance and, thus, only need to make a binary judgement about whether they match. In other words, there is less room for other factors to interfere with children's responses.

One weakness of TVJ data relates to the interpretation of negative responses. When a child says a statement is false after a context in which an adult would say it was true, it suggests the child did not access that interpretation of the sentence at that time. It cannot tell us, however, that the child is incapable of accessing that interpretation altogether. For example, a child might reject an interpretation consistently due to their lexical biases, but their parser is still capable of accessing that interpretation. Thus, claims about lack of competency based on TVJ data should be limited (Crain, 1998). One important consideration when working with TVJ data relates to statistical analyses. TVJ data is inherently binary, so the data is not normally distributed. This violates the assumptions of several statistical tests including linear models (Winter, 2013).

In sum, the strengths that made the TVJ task suited for Study 1's research questions were its simplicity for young subjects, its ability to rotate one utterance through multiple meaning contexts, and the fact that it does not require any metalinguistic awareness to perform the task. For the research questions in Study 1, these strengths outweighed the task's limitations such as its relatively slow pace and need for careful attention to pragmatic factors.

3.0 Study 2 – Variation on a picture selection task

Study 2 used a picture selection task to probe 4- to 7-year-olds' metalinguistic awareness of two types of syntactic ambiguity**.** In order to make the task engaging, the experiment was presented as a game in which children were charged with helping a puppet understand why some sentences mean one thing and some sentences mean two things. The puppet was presented as a "goofy" character who was forgetful and needed help learning about why a sentence can mean more than one thing. This had two purposes. First, the children were more comfortable explaining to a puppet than they likely would have been explaining to an adult (Thornton, 1996). Second, the puppet asked for each utterance to be repeated before the explanation phase in order to avoid taxing children's working memory. This repetition seems less odd to children because of the puppet's forgetful persona.

The picture selection task has been used extensively to test children's lexical and syntactic comprehension. An early and well-known example of the task comes from Fraser, Bellugi, and Brown (1963). They used picture selection to test children's understanding of certain morpho-syntactic contrasts (e.g., negative vs. affirmative, subject vs. object passive). It has been used frequently to probe disparities between comprehension and production by testing children's comprehension of constructions that they do not yet produce (Shipley, Smith, & Gleitman, 1969; Katz, Baker, & McNamara, 1974; Gerken & McIntosh, 1993, etc.). It has also been used to test children's comprehension of reflexives (Deutsch, Koster, & Koster, 1986) and more. More relevant to my research, Kidd and Bavin (2005) used the task to test the influence of lexical bias and definiteness on 5- to 9-year-olds' comprehension of syntactically ambiguous utterances.

The picture selection task has some important strengths. For example, like the TVJ task, it is easy for young children. In most picture selection tasks, children only need to point to images. This makes it more accessible to young and shy children than some other comprehension tasks, such as the act-out task (Gerken & Shady, 1996). The simplicity of the picture selection task was important for Study 2 as it made the task easy and fun to the children on the young end of the study age range. Previous studies used an interview task where children were expected to spontaneously notice an ambiguity (e.g., Wankoff, 1983; Cairns et al., 2004). Young children performed poorly at ambiguity detection in these studies, possibly because they were uncomfortable with the task. The picture selection portion of Study 2 gave children a very easy task to perform that felt more like games they are already familiar with (e.g., Memory). This likely made them more comfortable with the experimental setting than in previous studies. Similarly, having the children explain the ambiguity to a puppet was also meant to increase their comfort (Thornton, 1996).

There are also some weaknesses of the picture selection task. One weakness is that some grammatical constructions or properties are difficult to depict in a static image (Gerken & Shady, 1996). Thus, there are grammatical constructions that cannot be easily probed using a picture selection task. A second weakness is that the task depends on children's interpretation of images that were developed by adults. This can be problematic if the experimenter does not attend to the ways in which children's picture interpretations are different from those of adults'. Children as old as 10 do not always interpret images the same as adults (Berman, Friedman, Hamberger, & Snodgrass, 1989). In picture naming tasks, children often provide different names than adults do and have higher rates of naming errors (Cycowicz, Friedman, Rothstein, & Snodgrass, 1997). Many factors influence children's naming of pictures, including visual complexity, concept familiarity, number of possible names, etc. (Snodgrass & Vanderwart, 1980; Johnson & Clark, 1988; Johnson, 1992). Thus, it is important to pilot any experimental images extensively to ensure that they will be interpreted reliably by children of the age group of interest. I piloted the images for Study 2 with 17 3- to 4-year-olds, who were about one year younger than the participants in the study. During piloting, I asked children to name all of the objects in the pictures as well as the actions that were occurring. If children named the objects or actions in ways that differed from adult patterns, I asked if they would accept the names we provided for the objects and actions. If children did not accept our description of an image, it was adjusted, either by making certain properties more salient or by changing out an object altogether.

Another design issue to consider when developing a picture selection study is the delivery of the linguistic stimuli. In Study 2, an experimenter delivered all of the experimental utterances. One problem with this method is that humans cannot deliver an utterance with consistent prosody all of the time. Since we know that prosody can affect children's processing of syntactic ambiguity (Snedeker & Yuan, 2008), this is a weakness of the study. Other methods of delivering linguistic stimuli, such as recorded or synthesized speech, also have a weakness in that children are often less responsive to such stimuli because they find the situation confusing (Gerken & Shady, 1996). Since the primary goal of this study was to elicit children's explanations of ambiguity, it was more important to deliver the utterances in a way that made children comfortable than to make the prosody perfectly consistent. That being said, experimenters for Study 2 practiced delivery of the items for several hours to improve the consistency of their prosody. The target prosody was neutral with stress on neither the verb nor the local noun phrase. I will now discuss the strengths and weaknesses of the two data types produced by the picture selection task used in Study 2.

3.1 Comprehension and production data

This task produced two types of data: comprehension and production. The comprehension came from the picture selections; the production data came from children's explanations. I will begin by discussing the strengths and weaknesses of picture selection data and then move to the production data.

Like TVJ data, an important strength of picture selection data is that it is easy to code. In an act-out task or a production task, children can give responses that are unexpected or ambiguous. These present a coding challenge for the researcher. However, picture selection data is quite constrained. In fact, in studies where the child is choosing between two pictures, the data is binary. Picture selection data also has some weaknesses. One important weakness of picture selection data relates to limitations in interpretation. We infer that when a child chooses a picture, it represents their interpretation of the sentence. However, it is possible that a child chose that image for other reasons. In fact, in some picture selection studies, children choose between only two pictures. Thus, they have a 50% chance of choosing the right image by chance. This makes it difficult for researchers to distinguish children's grammatical knowledge from chance performance. Because of this issue, it is better to have children choose between more than two pictures if possible (Gerken & Shady, 1996). Study 2 faced an interpretation issue as well. Specifically, there were reasons a child could pick both target pictures for an ambiguous item other than ambiguity awareness. For example, the two target pictures were more similar to each other visually than the distractor pictures. A child could select the two target pictures based on this similarity rather than their relevance to the utterance. For example, if a child felt unsure if she was playing the game right, she might use a picture matching strategy because it is more similar to games she had previously played such as the card matching game Memory. Because of this, our interpretation of children's picture choices in Study 2 must be limited. This interpretation issue is why Study 2 also collected production data.

The main strength of production data is that it gives us a more direct window into children's grammatical abilities. If a child produces a grammatical construction, it provides strong evidence that the child has grammatical knowledge of that construction (Thornton, 1996). Similarly, in Study 2, children's explanations gave much stronger evidence of syntactic ambiguity awareness than their picture selection. If a child can explain why an utterance has two meanings, it indicates a level of conscious awareness of ambiguity. For example, the explanation in (7) was elicited from a 6-year-old in Study 2.

(7) "Because look it can be like he's pointing with an umbrella to the kid or he's pointing to the kid with an umbrella. Even though the pictures are different there's still an umbrella, a man, and a kid."

In this example, the child explicitly acknowledges the two interpretations of the utterance and discusses how those interpretations are similar and different. While it is possible that a child could select both pictures without a conscious awareness of ambiguity, it is unlikely that a child would give an explanation like in (7) if that child was not aware of the ambiguity.

One weakness of production data is that it often underrepresents children's grammatical knowledge. Because comprehension generally precedes production, children may be able to comprehend a construction well before they can produce it. Thus, when a child fails to produce a construction, it is impossible to know if it is simply a failure of the production system or a lack of underlying competency. Another weakness of production data is that it is more difficult for the participant. It can make shy or quiet children uncomfortable. Study 2 allayed this issue by including the easier picture selection portion first which made children more comfortable. A final weakness of production data is that it is not easily coded or, once coded, quantified. In order to have a numerical measure of children's ambiguity awareness in Study 2, I created a scoring system based on properties of adult explanations of their own picture choices. Specifically, explanations from an adult control group typically highlighted both the difference and similarity between the two interpretations of an ambiguous utterance. Thus, the following scale, which will be fully motivated in Appendix B, was developed.

Table 1: Explanation scoring system.

Score	Picture Choice	Participant Explanations
0	only one target or distractors	N/A
1	both targets	discussed choices without explicit similarity or difference
2	both targets	discussed either explicit difference or similarity
3	both targets	discussed both explicit difference and similarity

This scoring system allowed me to quantify explanations of syntactic ambiguity. However, it has its own weaknesses. It may underrepresent the number of children who noticed the ambiguity as some children may detect the ambiguity but not have the verbal skills to articulate it. Additionally, because it only focuses on two properties of adult explanations, it may miss other properties that indicate ambiguity awareness. For example, in Study 2, a feature of many adult explanations of adjective ambiguities was mention of the adjective's underspecificity. This property is not in the scoring system because it only applies to adjective ambiguities and the study also included PP ambiguities. There may be other construction-specific properties of adult explanations like this that my coding system does not measure. There is also a way in which this scale could over-represent children's awareness of ambiguity in that children could score highly by simply describing the differences and similarities between the images without actually noticing the ambiguity. From the researcher's perspective, we sometimes cannot be sure if the child is referring to the ambiguity per se or just the images.

To summarize, Study 2's adaptation of the picture selection task offers a new method for assessing children's awareness of the syntactic ambiguity. The picture selection portion of the task collects comprehension data and also engages children in the game. The explanation portion of the task provides production data and more direct evidence of ambiguity awareness. Collecting both data types together helps to overcome the weaknesses of each individual type.

4.0 Study 3 – Longitudinal training task

Finally, Study 3 used a longitudinal training design to show that 6- to 7-year-olds can improve their detection of syntactic ambiguity from regular practice activities. It also showed that improvement at ambiguity detection correlated with improvements on measures of reading readiness. The study lasted six weeks for each participant. All participants took a set of pre-tests in the first week and post-tests in the sixth week. These pre-/post-tests measured ambiguity awareness using the task from Study 2 and markers of emergent literacy using reading readiness tests. During the middle four weeks, children were divided into an ambiguity group and a math group. The ambiguity group took part in weekly sessions where they played games designed to improve their ability to detect and discuss syntactic ambiguity. This was the test group. The math group also took part in weekly sessions, but these children played games to improve their mathematics skills. This was the control group. The math control group was included because children may naturally improve at the ambiguity detection task and reading readiness tests over six weeks of attending school. By comparing children in the ambiguity group against those in the math group, we can determine if any improvements were a direct result of the ambiguity training.

Because the task used for the ambiguity detection pre-/post-test is the same as in Study 2, I will not discuss it further. Instead, I will focus on the strengths and weaknesses of a longitudinal design. The main strength of a longitudinal design is it allows the researcher to measure an individual's development or learning over time. Many studies have tracked children's language development over periods of years looking at the effect of factors such as early speech perception, cognitive impairment, home environment, and more (Elardo, Bradley, & Caldwell, 1977; Tsao, Liu, & Kuhl, 2004; Goldin-Meadow et al., 2014). Relevant to my research, Yuill (2009) and Zipke et al. (2009) used longitudinal designs to look at ambiguity detection learning in 7- to 9-year-olds by comparing reading scores in children who practiced ambiguity detection by discussing jokes and riddles. They found that there was a positive correlation between improvement at syntactic ambiguity detection and improved reading scores.

There are, however, certain weaknesses of longitudinal designs. Many of these relate to time and resources. First, because researchers must see the same subject multiple times over an extended period, there is a high risk of attrition. Specifically, for Study 3, if a child had missed more than a week because of illness or vacation, that child would have been removed from the study. Similarly, this type of research involves a much greater time commitment from research sites, making them difficult to recruit. Such designs also require a greater time commitment from the researchers themselves both for data collection and data analysis. Thus, more financial support is needed for an individual experiment than for many other types of language development studies.

4.1 Longitudinal data

One important consideration when working with longitudinal data is distinguishing correlation from causation. Some longitudinal designs can speak directly to causation. Others only measure correlation. Study 3 included a test group and a control group because it was designed to probe causation. Since the ambiguity detection training was the only systematic difference between the two groups and the control group did not improve at ambiguity detection, we can infer that the test group's improvements at ambiguity detection were caused by the training. Other types of longitudinal designs only probe correlations. For example, Cairns et al. (2004) measured syntactic ambiguity awareness in the same set of children in kindergarten and first grade. They also accessed the same children's standardized reading test scores in first and second grade. Thus, they were able to show that there was a correlation between children's ambiguity detection in first grade and their reading comprehension in second grade. This could not speak to causation though. From their data, there is no way to tell if syntactic ambiguity detection helps reading or the other way around.

Longitudinal data has its own particular set of strengths and weaknesses. The strengths of longitudinal data stem from the fact that it involves repeated measures from the same subject over time. This has several benefits. By analyzing an individuals' data over time, researchers can measure learning or development accounting for individual variation. Relatedly, longitudinal data allows for within-subject analyses, which inherently have more statistical power than between-subject analyses because the variability is reduced (Keppel & Wickens, 2004). Thus, one needs less subjects overall than in a between-subjects design. There are also some weaknesses of longitudinal data. First, data analyses can be a challenge for longitudinal studies because they typically collect more data overall than other types of studies. Furthermore, because a longitudinal design involves seeing the same subject several times, the likelihood is increased of having missing data points due to a subject missing one or more sessions. This issue can make statistical analyses more complex.

In sum, a longitudinal design was ideal for directly testing whether syntactic ambiguity training leads to improvement at the skill. By including a test group and a control group, it was able to speak to causation rather than just correlation. For Study 3, these advantages outweighed the fact that such a longitudinal design requires more time and resources than other studies.

5.0 Conclusion

To conclude, I have motivated the methods used in the four studies described in this dissertation's three journal articles (Appendices A-C) and Appendix 1. My dissertation research included judgement, comprehension, production, and reading readiness test data. The judgement data from Study 1 improved our understanding of children's comprehension of syntactic ambiguity by showing that children can access both interpretations of a syntactic ambiguity. The comprehension and production data from Study 2 gave nuanced measures of children's ambiguity detection that revealed earlier competency than previous research had. Comprehension and production data were also gathered in Study 3 along with the reading readiness test data to show that first graders can improve at syntactic ambiguity detection with practice and that improvement at syntactic ambiguity detection correlates with improvement on reading readiness measures. While each of these data types has its own strengths and weaknesses, the use of several types helps to overcome the weaknesses of each. In sum, by using several methods, this research project as a whole improves our understanding of children's syntactic ambiguity awareness and the comprehension that supports it. In the following chapter, I describe my overall conclusions from the results of these studies and some future research directions.

Chapter 3. Conclusions

In sum, the research described in this dissertation has advanced our understanding of children's syntactic ambiguity awareness in several ways. Study 1 showed that young children can access both interpretations of a syntactic ambiguity. Study 2 showed that some children's natural development of syntactic ambiguity awareness begins earlier than previous research had concluded. Most importantly, Study 3 showed that children younger than 8-years-old can improve their syntactic ambiguity awareness with practice and that such training correlates with improvement at reading readiness skills. In other words, syntactic ambiguity detection is a learnable skill that correlates with other skills known to aid in reading development. This suggests that syntactic ambiguity detection might usefully become part of literacy curricula and reading intervention strategies. Future research is needed to determine the precise nature of the relationship between syntactic ambiguity awareness and reading development. After I summarize the implications for the three studies, I will lay out several future research questions that would further our understanding of syntactic ambiguity awareness and its contribution to literacy development.

Study 1 improved our understanding of children's comprehension of syntactic ambiguity. It showed that 3- to 5-year-olds can access both interpretations of a syntactic ambiguity. This tells us that children's failure at previous ambiguity detection tasks is not due to an inability to access both interpretations of an ambiguity. This is important because we know that children's comprehension per se is not prohibiting children from noticing syntactic ambiguity in general. It still remains possible that linguistic factors could prevent children from detecting individual instances of syntactic ambiguity. As Cairns (2013) points out, if a child fails to revise a parse of an individual ambiguous sentence, the child would likely not notice both interpretations of the sentence. Study 1 did not test parsing, so the effect of online parsing on syntactic ambiguity detection is a promising avenue for future research. The results of Study 1 are also valuable because previous research on children's comprehension of ambiguity (e.g., Trueswell et al., 1999; Snedeker & Trueswell, 2004; Kidd & Bavin, 2005) has assumed that children could access all interpretations of an ambiguity, but until this study, it has never been directly tested. This study fills that gap and supports the suppositions of previous research. Furthermore, it demonstrates that children as young as 3-years-old can handle the added complexity involved in the comprehension of a syntactically ambiguous sentence.

Study 2 improved our understanding of 4- to 7-year-olds' metalinguistic awareness of syntactic ambiguity. It showed that some children have metalinguistic awareness of syntactic ambiguity earlier than previous research suggests (e.g. Cairns et al., 2004; Wankoff & Cairns, 2009). About 16% of these younger children showed a strong awareness of syntactic ambiguity. Many more children showed the beginning of syntactic ambiguity awareness development. For example, they chose both pictures for ambiguous items even though they did not explain the choice well. In fact, there was generally a considerable amount of variation in children's ability both on picture selections and explanations of the ambiguity. This suggests that syntactic ambiguity detection is a complex skill, and its component parts may develop at different times and rates for different children. It is important that we understand the development of syntactic ambiguity awareness because it likely support readings development. This is exemplified by Yuill (2009) and Zipke et al. (2009) who found that improvement at syntactic ambiguity awareness in 7- to 9-year-olds correlated with improved reading comprehension. Since Study 2 has revealed that some 5- to 6-year-olds are aware of syntactic ambiguity, it is likely that the connection to reading begins earlier in some children. If so, this would be meaningful for early literacy curricula. In particular, syntactic ambiguity awareness could be important for reading intervention strategies targeting children with poor comprehension but normal decoding skills.

Study 2 probably found more indications of early syntactic ambiguity awareness than previous research did because of key differences in the methods. This study used a variation on a picture selection task designed to test for conscious awareness of syntactic ambiguity. Children first chose from a set of four pictures all the ones that they thought matched the sentence. Then, children explained their choices to a confused puppet. This task had several benefits. The picture selection portion provided visual support for the ambiguity and also was engaging for children, while the puppet made children more comfortable teaching. Furthermore, having both picture selections and explanations provided a more nuanced measure for detecting early awareness of syntactic ambiguity. Children's picture selections alone provided weak evidence for their ambiguity detection. For children without the verbal skills to explain why a sentence is ambiguous, their picture choices revealed early signs of ambiguity awareness. Explanations provided stronger evidence for ambiguity awareness. The coding system developed for the study allowed for comparison of children's explanations to adult explanations. Thus, the study was able to measure gradient differences in children's ability to explain syntactic ambiguity and thus indicate a conscious awareness of the ambiguity.

Finally, Study 3 showed that practice can improve syntactic ambiguity detection in 6- to 7-year-olds. One hour of ambiguity awareness practice over four weeks improved children's ambiguity awareness scores by an average of about 4 points out of 12. This is especially notable in comparison to the children who practiced math activities and did not improve at all on the ambiguity awareness test. The two groups similar pre-test scores (i.e., 2.1 for the ambiguity group and 2.5 for the math group) suggests that it truly was the ambiguity awareness practice that caused the ambiguity group to improve and not underlying differences between groups. This is important because, if syntactic ambiguity does indeed aid in early reading development, then it matters that children can learn the skill through practice. Furthermore, the children who practiced syntactic ambiguity also improved more than the math group on the reading readiness tests. Again, the two groups had similar composite pre-test scores (i.e., 23.5 for the ambiguity group and 25.2 for the math group), indicating that they were evenly matched at the beginning of the study. This suggests that improvement at ambiguity detection may correlate with improvement on other metalinguistic skills that we know support reading development. This indirectly supports the conclusions of previous research that syntactic ambiguity detection contributes to early reading development (e.g. Yuill, 2009; Cairns, 2013). Thus, my findings add to a growing body of literature suggesting that syntactic ambiguity awareness can and should be part of early reading curricula.

However, before syntactic ambiguity detection can be used in curricula and intervention strategies, more research is needed to understand the specific details of its development and its role in reading. One future avenue of research should investigate other factors that affect the development of syntactic ambiguity detection, such as socioeconomic background. A more detailed examination of the factors that correlate with ambiguity awareness in 5- to 7-year-olds could elucidate why Study 2 revealed such variation across children. For example, there are correlations between a child's metalinguistic awareness and the cognitive ability of conservation (Hakes, 1980; Schlisselberg, 1988; van Kleeck, 1994). Another factor that might contribute to children's awareness of syntactic ambiguity is each child's language background. There is considerable evidence that certain metalinguistic skills are advanced in multilinguals (e.g., Bialystok, 1986; Davidson et al., 2010; Hermanto et al., 2012). However, there are several social factors that may influence the degree to which bilinguals are metalinguistically advanced. For example, children who translate for their parents (termed "child brokers") tend to have particularly advanced metalinguistic skills (Raschke, 2013). Balanced bilinguals have better metalinguistic skills than unbalanced bilinguals (Bialystok & Barac, 2012). There is also research showing that bilinguals who are literate in both languages have more advanced metalinguistic skills than bilinguals who are literate in only one (Thomas, 1988). Relatedly, research also suggests that the discourse styles of a child's family and community could affect his metalinguistic development (Heath, 1983). Another cognitive factor worth exploring might be the relative prestige level of the bilingual child's two languages. Thus, language background should be probed in future research on children's ambiguity awareness. By isolating the factors that contribute to children's syntactic ambiguity awareness, we can better understand how to incorporate it into reading curricula meant for diverse children.

It would also be of interest to teach children syntactic ambiguity awareness with more types of syntactic ambiguity. My own intuitions and those of my colleagues suggest that certain types of syntactic ambiguity are easier to detect than others. It is unclear whether children would find some structural ambiguities more difficult to detect or learn to detect. Relevant examples include bi-clausal attachment ambiguity (as in *Where did Mary hear the man singing?*), adverbial attachment ambiguity (as in *Children who sing often dance.*), and ellipsis ambiguity (as in *Susan asked Bill to clean up, and Mary too.*). Such ambiguities may be a more difficult ambiguity to detect than a prepositional phrase ambiguity or adjective ambiguity. Thus, training children to detect such an ambiguity might benefit their metalinguistic skills more.

Future research should also address the question of training even younger children to detect syntactic ambiguity. Study 3 focused on first graders, but it may be that preschoolers can learn such skills with direct instruction. This matters because we know that children from certain socioeconomic groups fall behind their peers in reading as early as preschool (e.g., Vanneman et al., 2009; Hemphill & Vanneman, 2011; Hoff, 2013) and that interventions are more successful the earlier they are applied (Vellutino et al., 1998). Furthermore, Study 3 showed that improvement at syntactic ambiguity awareness correlates with improvement on reading readiness measures. These measures are correlated with reading but are not reading itself. Thus, a study directly measuring improvement at reading in first graders would be valuable. One possibility would be a training study similar to Study 3 but with standardized reading tests as pre-/post-tests. One problem with this line of reasoning is that some first graders are not yet reading well enough to participate in such a test. Thus, one would either need to eliminate the poorest readers, creating a sampling bias, or one would have to measure some children with a reading test and some with reading readiness tests, which is also problematic. Another possibility would be a longitudinal design similar to Cairns et al. (2004), where ambiguity awareness is measured in kindergarten and first grade and then tested for correlations with second grade standardized reading scores. The problem with this type of design is that the relationship can only be correlational. There is no evidence for causation. Training studies such as Study 3, on the other hand, suggest direct causation.

A final future research direction relates to a possible criticism of Study 3. Specifically, it is possible that the ambiguity groups' improvement on the reading readiness tests was not due to improvement on syntactic ambiguity detection per se. Rather, it might have been due to the verbal nature of the ambiguity games. In other words, it might be improvement on general verbal skills correlating with reading readiness measures and not specifically ambiguity detection. One way to tease this apart would be to conduct a similar study with a control group that does verbal activities rather than math activities. This group would participate in activities with a focus on verbal skills that are not directly related to syntactic ambiguity detection, such as vocabulary development, passage reading, and story development. Like in Study 3, the test group would participate in activities to practice syntactic ambiguity detection. Thus, if the test group who practiced ambiguity detection still improved more on the reading readiness measures than the control group who practiced other verbal skills, it would suggest that syntactic ambiguity awareness truly does contribute independently to reading development.

Appendix 1: Lexical Biases in Young Children's Ambiguity Resolution

1.0 Introduction

This appendix describes a pilot study conducted in the Speech, Language and Brain Laboratory with Dr. Gayle Dede. This goal of the study was to probe how verb biases influence young children's processing of prepositional phrase (PP) ambiguities. Previous research on this question used an act-out task, which is relatively slow. Our intention was to use eye-tracking, to gather online data on young children's ambiguity resolution. We hypothesized that the effect of lexical bias on online sentence processing would be weaker for 2- to 3-year-olds than for 4-year-olds because the younger children are still learning certain co-occurrence biases. This hypothesis is supported by Kidd and Bavin (2005) who found that 3-year-olds' performance in an act-out task was less influenced by verb bias than that of 5-year-olds. The younger children tended to act out the VP-interpretation of the sentence regardless of verb bias. This suggests that children's verb biases in regard to PP ambiguities may develop over time. However, an act-out task with no online data can only provide weak evidence for any parsing claims. It is possible that the 3-year-olds performed differently than 5-year-olds because of other factors. For example, the sentences in the study were commands such as *Cut the cake with the candle.* Younger children may have found acting out the VP-interpretation, which often involved using unusual instruments, to be more enjoyable. Thus, online data is needed to determine whether younger children truly have weaker verb biases. No study has ever collected such data to probe syntactic ambiguity processing in children younger than 4-years-old.

Previous research has found a robust effect of verb bias in adults' and older children's resolution of PP ambiguities though (Spivey-Knowlton & Sedivy, 1995; Snedeker & Trueswell, 2004). These studies have used two types of methods: self-paced reading and an act-out task paired with eye-tracking. These methods were not ideal for our study on very young children's processing of such ambiguities. Self-paced reading would not work because 2- to 4-year-olds are not yet reading. An act-out task would not work because of the type of eye tracker we used, which tracks eye movements on a computer screen. The act-out task was also not ideal because children as young as 2-years-old are easily distracted by the toys and props in such a task (Goodluck, 1996). Because of this, we chose a passive listening task where images on the computer screen corresponded to different interpretations of the ambiguous sentence. Participants looked at the images while listening to a sentence. Then, they answered questions from a puppet that were unrelated to the ambiguity. These questions were intended to keep the participants' interested in the task.

Results from a pilot group of 10 adults did not show any trends in regards to verb bias. As the effect of verb bias on adult processing has been replicated in several studies, our failure to find any such effect suggests the passive looking paradigm in our experiment was not appropriate for eliciting such verb biases in participants. Because of this, we did not test any child participants. The following report summarizes the rationale, methods, and results of the pilot study, along with a discussion of why it may not have worked.

2.0 Background literature

Many years of research have shown that lexical co-occurrence biases affect ambiguity resolution (Spivey-Knowlton & Sedivy, 1995). These biases cause adults and children to have preferences for one interpretation of an ambiguous utterance over another based on the semantic category of the verb (Spivey-Knowlton & Sedivy, 1995; Snedeker & Trueswell, 2004; Kidd & Bavin, 2005). Specifically, action verbs such as *poke* or *hit* are biased towards VP-attachment, or an instrument interpretation. Perception verbs such as *see* or *hear* are biased towards NP-attachment, or a modifier interpretation (Spivey-Knowlton & Sedivy, 1995). However, Kidd and Bavin (2005) showed that 3-year-old's comprehension of ambiguous sentence was less affected by lexical biases than older children's.

Lexical co-occurrence biases are just one of many informational cues that adults use during ambiguity resolution. Other cues include animacy of participants (Trueswell, Tanenhaus, & Garnsey, 1994), referential information (Altmann & Steedman, 1988), and probabilistic verb biases (Spivey-Knowlton & Sedivy, 1995). Children seem to use fewer informational types when processing ambiguities. For example, Trueswell et al. (1999), which used an act-out task monitored by eye-tracking to examine 5-year-old children's processing of temporary ambiguities, found that adults' eye gaze patterns and responses were influenced by number of referents in the visual world context. Children's eye gaze patterns and responses were the same regardless of the number of referents.

Snedeker and Trueswell (2004), as well as Kidd and Bavin (2005), provided further support that children do not attend to number of referents while processing syntactic ambiguities. Instead, they argued that children rely primarily on probabilistic verb biases when resolving PP ambiguities. Snedeker and Trueswell (2004) paired eye-tracking with an act-out task to test whether 5-year-olds were sensitive to these verb biases and, if so, whether they pattern by the same semantic categories as adults. The study had three verb bias conditions: verbs with NP-attachment bias, verbs with a VP-attachment bias, and neutral verbs. They found that both children and adults' gazes were influenced by lexical bias in the sense that they would look earliest and longest at the toy that matched their interpretation bias. However, only the adult group seemed to also take into account other informational types such as number of referents. Kidd and Bavin (2005) corroborated this finding. They used a picture selection task to test whether 5- to 9-year-olds demonstrated lexical biases consistent with those previously described. They found that children's selections did correspond to adult biases.

Kidd and Bavin (2005) also suggest a developmental trend with respect to verb biases. In a second experiment, they used an act-out task to compare two age groups (3- & 5- year-olds) and their comprehension of syntactic ambiguities. Their participants acted out sentences that were manipulated for lexical bias and plausibility. One such sentence is exemplified in (8).

(8) Cut the tree with the leaves.

In this case, the verb in the sentence is an action verb and thus biased towards VP-attachment. However, the VP-attachment reading is highly implausible because it involves using leaves as an instrument for cutting a tree. All of the sentences that children acted out were manipulated so that they varied by verb bias and whether or not the verb bias supported a plausible reading of the sentence. They found that most 3-year-olds and some 5-year-olds ignore plausibility when interpreting sentences with PP ambiguities, further supporting the notion that children use less informational types than adults do during sentence processing. Interestingly, they also found that the 3-year-olds seemed to have less clear lexical biases than the 5-year-olds. Rather, the 3-year-olds seemed to have a general preference for the instrumental interpretation regardless of the lexical bias of the verb. Kidd and Bavin argue that this may be because the 3-year-olds are still learning the lexical co-occurrence restrictions associated with verbs that have an NP-attachment bias. This suggests a developmental trend in which older children have more established lexical biases than younger children.

To summarize, this research shows that children do not apply the same informational cues as adults do while processing ambiguities. They seem to rely heavily on lexical co-occurrence biases, while adults can incorporate other informational types. Because of children's strong reliance on this cue, it is important to understand its developmental trajectory. Kidd and Bavin (2005) suggest that the younger the child is, the less established their lexical biases may be. This makes sense given that verb biases are formed by attending to lexical frequency information, and younger children have had less access to such information than older children and adults. However, no study has used an online measure to test the effect lexical bias on ambiguity resolution in children younger than 4-year-old. It is important to do so because only online data can give real-time evidence about the influence of verb biases on children's processing. Our study was intended to fill that gap. Because the results of the adult pilot study indicated that our task was not ideal for revealing the effect of verb biases, we did not conduct the study with children. We report below the methods and results for the pilot study. We also discuss some reasons that it may not have worked with recommendations for future research.

3.0 Methods

In this section, we describe the methods of the pilot study including details about the participants, materials, and task.

3.1 Participants

Ten adults participated in the pilot study. They were undergraduate students at the University of Arizona.

3.2 Materials

All the experimental items were prepositional phrase ambiguities said in a neutral prosody. Participants heard 30 experimental items and 16 filler items, shown below in Table 2. Filler items were simple sentences of varying structures. Ten of the experimental items contained verbs biased towards VP-attachment, 10 contained verbs biased towards NP-attachment, and 10 contained verbs with no bias. Verbs were taken from Snedeker and Trueswell (2004) who normed them using a sentence completion task.

Table 2: Experimental and filler items.

Instrument Bias	
1. The lady tickled the baby with a feather.	2. The man poked the monkey with a banana.
3. The girl hit the man with a flower.	4. The man fed the baby with a spoon.
5. The lady bopped the puppy with a ball.	6. The lady covered the kitten with a blanket.
7. The lady touched the owl with a flower.	8. The boy brushed with the horse the comb.
9. The boy whacked the man with a balloon.	10. The man cleaned the bunny with a towel.
Modifier Bias	
11. The boy looked at the girl with binoculars.	12. The man found the girl with a map.
13. The woman tasted the soup with a spoon.	14. The woman sang to the girl with a microphone.
15. The lady listened to the boy with a tube.	16. The man listened to the girl with the walkie talkie.
17. The man hugged the baby with a blanket.	18. The boy yelled at the girl with a megaphone.
19. The girl heard the man with headphones.	20. The boy watched the girl with a magnifying glass.
Neutral Bias	
21. The boy painted the girl with a brush.	22. The man dragged the girl with a wagon.
23. The lady felt the hamster with a ribbon.	24. The lady threw the mashed potatoes with a fork.
25. The girl scratched the kitten with a comb.	26. The man turned over the turtle with a stick.
27. The girl blew on the teddy bear with a fan.	28. The woman drew the man with a crayon.
29. The boy pinched the lady with a barrette.	30. The man pointed to the girl with an umbrella.
Fillers	
31. The man chased the dog with the leash.	32. The girl wanted the doll that has yellow hair.
33. The boy surprised the girl with a scarf.	34. The woman kicked the ball that has stars
35. The man used a pen to write a letter.	36. The lady sang to the puppy with a microphone.
37. The lady broke the red bowl.	38. The boy helped the girl who's wearing pants.
39. The girl filled the jar with marbles.	40. The lady used a spoon to eat her lunch.
41. The man held the book with a sock.	42. The woman stared at the hamster with a carrot.
43. The woman lifted the guitar with a strap.	44. The lady used a stick to move the tricycle.
45. The girl bent the flower with a book.	46. The boy used a doll to push the boat.

We created the pictures on Microsoft Paint. Pictures were normed for visual interest. Ten adults silently looked at each set of pictures. All items containing pictures that attracted an unusually large or small percentage of looking time were removed before the study to avoid items with images that were unbalanced in terms of visual interest.

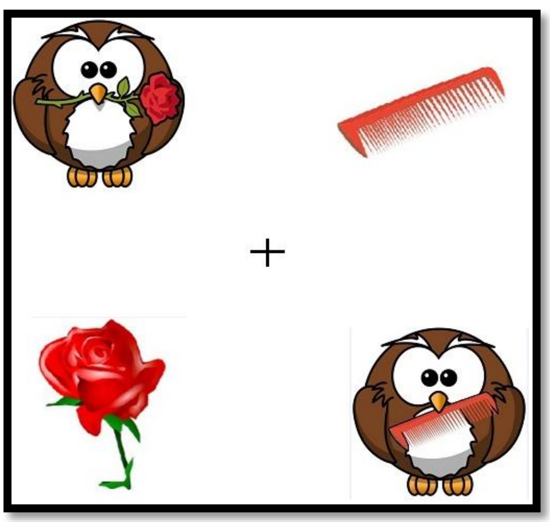
3.3 Task

The study paired eye-tracking with a visual world paradigm. Participants were seated in front of the eye-tracker. Their eye movements were recorded while they looked at sets of four pictures and heard PP ambiguities. For each sentence, participants saw four pictures on the screen: the instrument of the sentence (target instrument), a distractor instrument, the target patient modified by the target instrument (target modified), and a distractor target with the distractor instrument (distractor modified). The four pictures were randomly located in the four corners of the screen. Take for example the sentence in (9).

(9) The lady touched the owl with the flower.

When participants heard this sentence, they saw a flower (target instrument), a comb (distractor instrument), an owl with a flower (target modified patient), and an owl with a comb (distractor modified patient), as shown in Figure 7.

Figure 7: *The lady touched the owl with the flower.*

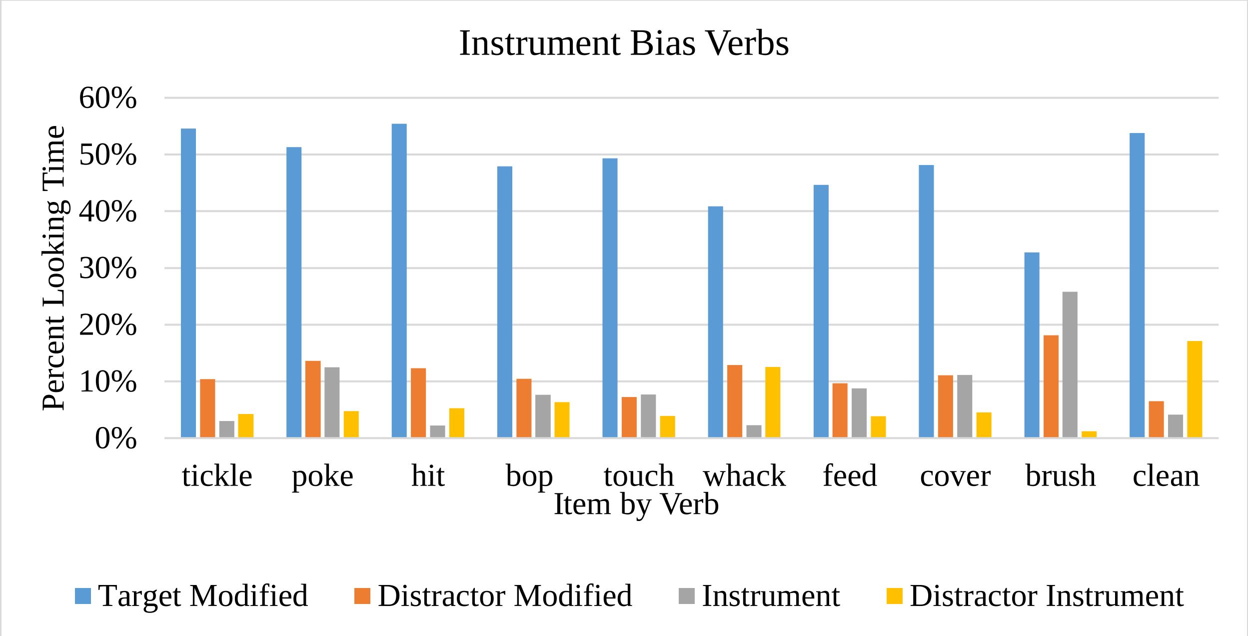


Participants were given three seconds to look at all pictures before they heard the sentence. They then heard the pre-recorded experimental sentence uttered by a female voice. After the ambiguous experimental utterance, participants heard a question uttered by a puppet. These questions were intended to keep participants engaged in the experiment but were unrelated to the experimental items. For example, questions asked about properties of the picture (e.g., *What color is the flower?*) and properties of the scene (e.g., *How many pictures of owls are there?*). Response to the questions were not recorded. After responding, participants saw a colorful and interesting picture as a reward.

4.0 Results

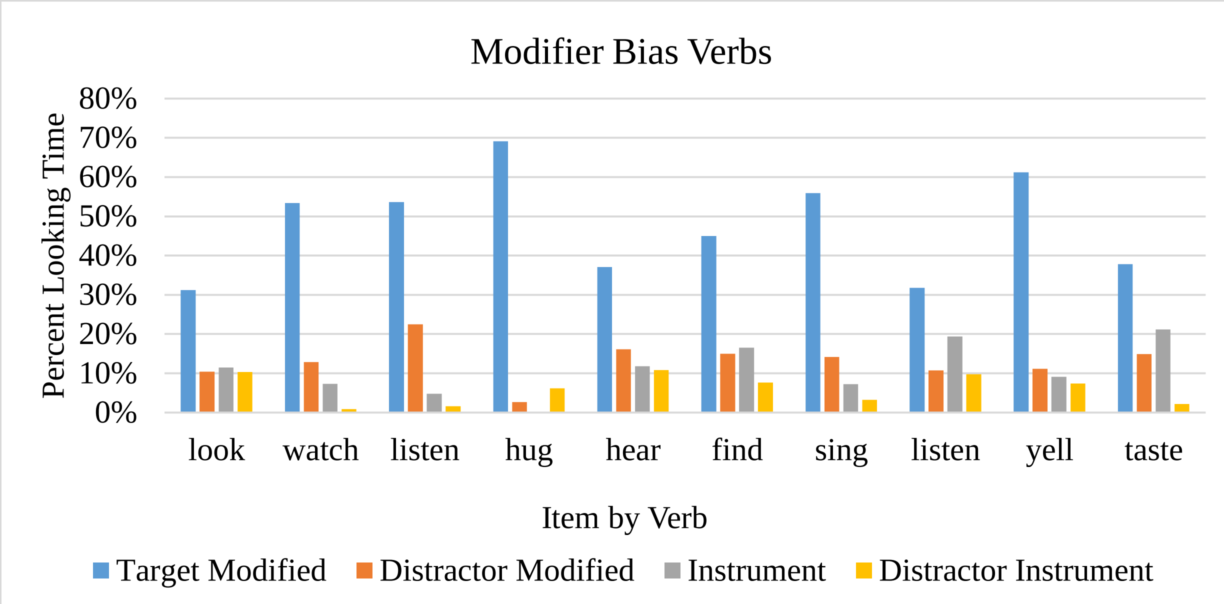
Participants' looking time at the pictures during and after the ambiguous experimental items were recorded beginning at end of the verb. The 10 adult pilot subjects showed no effect of verb bias. Results showed that participants looked at the target modified image significantly more than the others, regardless of verb bias. Overall, participants look most at the target modified picture on 28/30 utterances. The modified distractor picture was the second most looked at image for 18/30 utterances. This is of particular note for the items in Figure 8 that should be biased towards the instrument interpretation. For these items, we expected participants to look most at the target instrument image. However, all of the items in this category elicited considerable more looking time at the target modified picture than the target instrument picture.

Figure 8: Percent looking time after the verb for sentences with instrument biased verbs.



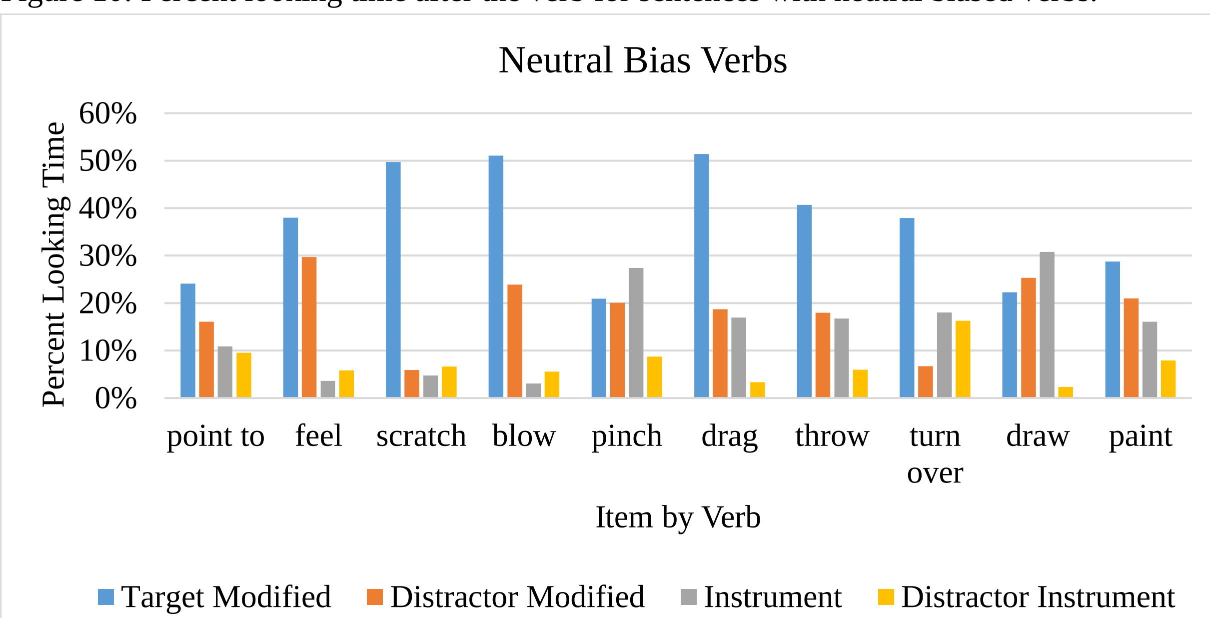
At the very least, we expected participants to look at the target instrument picture more for the instrument-biased items in Figure 8 than they did for the modifier-biased items in Figure 9. However, the instrument-biased items generally elicited about the same or sometimes more looking time at the target modified picture than the modifier-biased items. For example, in Figure 9, modifier-biased utterances with the verbs *look, listen, taste,* and *hear* all elicited a smaller proportion of the looking time at the target modified picture than the 9/10 instrument-biased verbs. This is interesting because these four verbs all fit solidly in the semantic category of perception verbs which past research posits is biased towards a modifier interpretation.

Figure 9: Percent looking time after the verb for sentences with modifier biased verbs.



The neutrally biased verbs were also surprising. Interestingly, two of the items with a neutral bias were the only items that showed a preference for the Target Instrument image as shown in Figure 10. Most of the neutral items showed a preference for the Target Modified Patient however. Generally, we expected that the neutral verbs would elicit closer to equal looking time between the Target Modified Patient and the Target Instrument than the other two bias categories. Rather, this category just showed a great deal of individual variation by item.

Figure 10: Percent looking time after the verb for sentences with neutral biased verbs.



5.0 Discussion

In sum, these results did not pattern as would be expected if participants' looking times had been influenced by lexical biases. Rather, participants had a general preference to look at the modified pictures more than the instrument pictures. Because of this, we concluded that the task was not right for this particular research question and did not to test any child subjects.

We would like to discuss why this task may have failed to find a well-established effect in the sentence processing literature. There are certain key differences between our task and those used in studies on lexical bias in children's processing. Specifically, ours was a passive looking task, while others have used an act-out task. Thus, the act-out task requires the participant to listen to an utterance and plan out their action (Saddy, 1992; Goodluck, 1996), while ours did not. As Trueswell and Tanenhaus (2005) point out, the act-out task, by forcing the participant to plan an action, may cause them to look more at each individual object as a possible object that might need to be moved. This planning process also may bias them to look more at the instrument objects than they otherwise would. It seems reasonable that a person planning to perform an action would look at the items in an array that are more likely to be instruments based on real world knowledge.

Our study did not require any action planning. Thus, participants' looking times may have been more based on visual interest than searching for a plausible instrument. This fits with our data that showed participants looking most at the two modified pictures, which were visually more complex than the two instrument pictures. Specifically, the modifier pictures were a person or animal and an object, whereas the instrument pictures were a single object. One way to avoid this issue in future research would be to control more precisely for visual complexity. Typically, this is done through a norming study in which participants rank the complexity of an image (Cycowicz et al., 1997; Székely & Bates, 2000). With images controlled for visual complexity, it would be interesting to probe whether a passive looking task finds an effect of verb bias. If it did not, this would raise questions about results obtained from studies with act-out tasks. Were results from instrument biased conditions inflated by the influence of the participant planning their own action? At the very least, this would suggest that verb biases may be somewhat task dependent. To tease apart these issues, it would be interesting to directly compare children's responses in a passive looking task and an act-out task with matched materials.

Appendix 2: Maintaining Relationships with Research Sites

When working with a child population, it can be difficult to find and recruit subjects. If a researcher's target population is in school, one way to reach many children is to conduct the research directly at schools. This is ideal from the researcher's perspective because participants can be recruited and tested in one place. However, from the school's perspective, hosting a research study causes a certain amount of disruption. For example, school directors and teachers must take time to schedule with the researchers. Often, it also requires that the school have an empty room for the research to take place in. For example, in Studies 1, 2, and 3 of this dissertation, I pulled individual children out of their classrooms to participate in the research. This meant that I had to coordinate with directors and teachers to schedule appropriate times for this. In some cases, teachers generously modified their lesson plans based on when the research was taking place. Study 3 was particularly demanding because it required that each child be removed from class once a week for six weeks. Thus, those teachers of participating classrooms who adjusted their lesson plans for the times when the research was occurring had to do so for a month and a half. Schools that are willing to accept these inconveniences do so because they value research and appreciate that it can help improve our understanding of child development and learning. Nevertheless, it is paramount that researchers respect the sacrifice the schools are making to support our research. I will discuss a few ways that researchers can minimize the inconvenience to schools and also give back to the schools to thank them for their partnership using specific examples from my dissertation research.

When recruiting and working with schools, a primary goal of the researcher should be to respect the wishes of the school's administration and teachers. This means basing your schedule on the school's preferences. For example, because Study 3 required children to be taken out of the classroom multiple times, several teachers requested that I do my research during already scheduled free work times. They would often request that we not come to the school during core subject times like math and reading. This made scheduling a greater challenge on our end, particularly since the research required two experimenters. Thus, I hired three undergraduate assistants so that there would be enough free time in their combined schedules to allow us to be accommodate teachers' scheduling preferences. This was possible because Studies 2 and 3 of this dissertation were supported by a Doctoral Dissertation Research Improvement Grant from the National Science Foundation (PI: Cecile McKee; BCS-1451665). Similarly, it is also important to coordinate with teachers regarding individual children. Often teachers will have preferences about when each child comes out of the classroom. For example, one teacher allowed me to take children who were very good at math during math lessons, but not children who struggled with math. Respecting teacher's choices in this regard shows them that you both appreciate their professional expertise and value the child's education over your own research agenda. There are other ways to minimize inconvenience for teachers. For example, to get permission from parents, we send home permission forms that the parents then return. I offered to the schools to put the forms in children's take-home folders myself, so that my research procedures did not burden teachers with extra work.

Another important piece of working with community partners is finding ways to thank them for their help. One way to do this is to report back to the schools on the findings of the study. As I mentioned earlier, many schools participate in research such as mine because they understand its value. This means that school staff are generally interested in knowing the results of the research that they facilitated. This may sound obvious, but multiple school directors have told me that researchers who worked at their schools previously never contacted them again after data collection was complete. Following up after the research is over and informing schools of the results is respectful and reflects more positively on the research community. It also makes schools more willing to host research in the future.

I have reported to schools on my research findings in a number of ways, depending on the preferences of the directors and other staff. For some schools, I have presented my studies and results at staff meetings or teacher training events. I try to make these presentations maximally useful to teachers, by focusing on applications of the results. I also make these presentations interactive by demonstrating parts of the experiment with teacher volunteers. Teachers found it amusing when we pulled out puppets and pictures. Teachers comments and questions after such presentations indicated that they were engaged and interested. For example, teachers at one school asked questions about children's responses to experimental items and practical applications of the research. One teacher commented on how much her students enjoyed participating in the study. This type of feedback indicates that we have developed a positive working relationship with that school. For other schools, I have developed accessibly written reports of the experiment design, results, and conclusions. This is preferable to some schools that do not have the flexibility to schedule a presentation for the staff. The written reports present the basic background research, methods, and results of the study with an emphasis on real world applications. It is also important to thank the parents who give permission for their children to participate in the research. Following a long standing tradition of the Developmental Psycholinguistics Lab at the University of Arizona, I wrote a short thank-you letter for parents at the end of the study. This letter described the status and goals of the research and thanked parents for supporting it.

Because Study 3 required more time and resources from the schools, I developed additional materials to thank them. For example, Study 3 includes several language and math learning games. Children in the study generally enjoyed these games a lot, so we made simple versions of some of the games for each classroom that participated. The versions of the games we gave to classrooms included simple instructions so that children could play them with very little help from the teachers. Teachers were surprised and pleased to receive these games. Something else I did for schools that participated in Study 3 was to give them flash drives with digital materials about the study. This included articles relevant to the research, a digital copy of the written report, and some of the images we used to teach about syntactic ambiguity that teachers can print to develop their own classroom activities.

Of course, making such materials requires funding. We included funds to create materials for schools in the budget of our NSF grant that supported this research. I think this should be part of any funding for research that will take place at schools or other sites in the community. Reporting the results of our research back to the research site is not only a way to maintain good working relationships with our community partners, it is also an important way to disseminate our findings to people for whom it may have a practical application.

References for Chapters 1-3 and Appendices 1-2

Abrams, K. H., Chiarello, C., Cress, K., Green, S., & Ellelt, N. (1978). The relation between mother-to-child speech and word order comprehension strategies in children. In R.N. Campbell & P.T. Smith (Eds.), *Recent advances in the psychology of language*. New York, NY: Plenum Press.

Adlof, S. M., Catts, H. W., & Little, T. D. (2006). Should the simple view of reading include a fluency component?. *Reading and Writing*, *19*(9), 933-958.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–238.

Andere, E. (2014). *Teachers' perspectives on Finnish school education: Creating learning environments*. Cham: Springer.

Aronsson, K. (1981). The bilingual preschooler as grammarian: Children's paraphrases of ungrammatical sentences. *Psychological Research Bulletin.* University of Lund.

Baumann, J.F., Hoffman, J.V., Moon, J., & Duffy-Hester, A.M. (1998). Where are teachers' voices in the phonics/whole language debate? Results from a survey Of U.S. elementary classroom teachers. *The Reading Teacher, 51*, 636-650.

Bentin, S., Deutsch, A., & Liberman, I. Y. (1990). Syntactic competence and reading ability in children. *Journal of Experimental Child Psychology*, *49*(1), 147-172.

Berman, S., Friedman, D., Hamberger, M., & Snodgrass, J. G. (1989). Developmental picture norms: Relationships between name agreement, familiarity, and visual complexity for child and adult ratings of two sets of line drawings. *Behavior Research Methods, Instruments, & Computers*, *21*(3), 371-382.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development*, *57*(2), 498-510.

Bialystok, E. (1992). Selective attention in cognitive processing: The bilingual edge. *Advances in Psychology*, *83*, 501-513.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70,* 636-644.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. New York, NY: Cambridge University Press.

Bialystok, E. (2001). Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics, 21*, 169-181

Bialystok, E., & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition*, *122*(1), 67-73.

Bialystok, E., & Ryan, E. B. (1985). Toward a definition of metalinguistic skill. *Merrill-Palmer Quarterly, 31*(3), 229-251.

Birdsong, D. (1989). *Metalinguistic performance and interlinguistic competence*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Cain, K., Oakhill, J., & Lemmon, K. (2005). The relation between children's reading comprehension level and their comprehension of idioms. *Journal of experimental child psychology*, *90*(1), 65-87.

Cairns, H.S. (2013). Metalinguistic skills of children. In M. Becker, J. Grinstead, & J. Rothman (Eds.), *Generative Linguistics and Acquisition: Studies in Honor of Nina M. Hyams*. Philadelphia, PA: John Benjamins Publishing Company.

Cairns, H. S., Waltzman, D., & Schlisselberg, G. (2004). Detecting the ambiguity of sentences: Relationship to early reading skill. *Communication Disorders Quarterly, 25*, 68-78.

Cairns, H. S., Schlisselberg, G., Waltzman, D., & McDaniel, D. (2006). Development of a metalinguistic skill judging the grammaticality of sentences. *Communication Disorders Quarterly, 27*(4), 213-220.

Carducci-Bolchazy, M. (1978). A survey of the use of reading readiness tests. *Reading Horizons, 18*(3), 10.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read?. *Cognition*, *91*(1), 77-111.

Cazden, C. B. (1974). Play and metalinguistic awareness: One dimension of language experience. *The Urban Review, 7*(1), 28-39.

Chall, J., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, Mass: Harvard University Press.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*(3), 687.

Crain, S., 1947, & Thornton, R. (1998). *Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics*. Cambridge, MA: MIT Press.

Crain, S., Thornton, R., Boster, C., Conway, L., Lillo-Martin, D., & Woodams, E. (1996). Quantification without qualification. *Language Acquisition, 5*(2), 83-153.

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

Davidson, D., Raschke, V. R., & Pervez, J. (2010). Syntactic awareness in young monolingual and bilingual (Urdu–English) children. *Cognitive Development*, *25*(2), 166-182.

Demont, E., & Gombert, J. E. (1996). Phonological awareness as a predictor of recoding skills and syntactic awareness as a predictor of comprehension skills. *British Journal of Educational Psychology, 66*(3), 315-332.

Deutsch, W., Koster, C., & Koster, J. (1986). What can we learn from children's errors in understanding anaphora?. *Linguistics*, *24*(1), 203-226.

Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school*. Baltimore, Md: P.H. Brookes Pub. Co.

Downing, J. (1979). *Reading and reasoning*. New York, NY: Springer-Verlag.

Duncan, G. J., & Sojourner, A. J. (2013). Can intensive early childhood intervention programs eliminate income-based cognitive and achievement gaps?. *Journal of Human Resources*, *48*(4), 945-968.

Ehri, L. C. (1989). The development of spelling knowledge and its role in reading acquisition and reading disability. *Journal of Learning Disabilities*, *22*, 356–365.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*(3), 393-447.

Elardo, R., Bradley, R., & Caldwell, B. M. (1977). A longitudinal study of the relation of infants' home environments to language development at age three. *Child Development*, *48*(2), 595-603.

Elkind, D. (1987). *Miseducation: Preschoolers at risk*. New York: Knopf.

Elkind, D. (2001). *The hurried child: Growing up too fast too soon* (3rd ed.). Cambridge, Mass: Perseus Pub.

Felser, C., Marinis, T., & Clahsen, H. (2003). Children's processing of ambiguous sentences: A study of relative clause attachment. *Language Acquisition*, *11*(3), 127-163.

Fraser, C., Bellugi, U., & Brown, R. (1963). Control of grammar in imitation, comprehension, and production. *Journal of Verbal Learning and Verbal Behavior*, *2*(2), 121-135.

Gathercole, S. E., Tiffany, C., Briscoe, J., & Thorn, A. (2005). Developmental consequences of poor phonological short-term memory function in childhood: A longitudinal study. *Journal of Child Psychology and Psychiatry*, *46*(6), 598-611.

Gerken, L., & McIntosh, B. J. (1993). Interplay of function morphemes and prosody in early language. *Developmental Psychology*, *29*(3), 448.

Gerken, L., & Shady, M. E. (1996). The picture selection task. In D. McDaniel, C. McKee, & H.S. Cairns (Eds.). *Methods for assessing children's syntax*, Cambridge, MA: MIT Press.

Gleitman, L. R., & Gleitman, H. (1971). *Phrase and paraphrase: Some innovative uses of language*. New York, NY: Norton.

Goldin-Meadow, S., Levine, S. C., Hedges, L. V., Huttenlocher, J., Raudenbush, S. W., & Small, S. L. (2014). New evidence about language and cognitive development based on a longitudinal study: Hypotheses for intervention. *American Psychologist*, *69*(6), 588.

Goodluck, H. (1996). The act-out task. In D. McDaniel, C. McKee & H. Cairns (Eds.), *Methods for assessing children's syntax.* Cambridge, MA: MIT Press.

Gordon, P. (1996). The truth-value judgment task. In D. McDaniel, C. McKee & H. Cairns (Eds.), *Methods for assessing children's syntax*. Cambridge, MA: MIT Press.

Gordon, P., & Chafetz, J. (1986). *Lexical learning and generalization in the passive acquisition*. Paper presented at the 11th Annual Boston University Conference on Language Development, Boston, MA.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hove: Lawrence Erlbaum.

Gough, P. B., & Tunmer, W. E. (1986) Decoding, reading and reading disability. *Remedial and Special Education*, *7*(1), 6–10.

Graves, M. F., & Slater, W. H. (1987). *The development of reading vocabularies in rural disadvantaged students, inner-city disadvantaged students, and middle-class suburban students*. Paper presented at the meeting of the American Educational Research Association, Washington, DC.

Grigorenko, E. L., Klin, A., & Volkmar, F. (2003). Hyperlexia: disability or superability?. *Journal of Child Psychology and Psychiatry*, *44*(8), 1079-1091.

Hakes, D. T. (1980). *The development of metalinguistic abilities in children*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. Cambridge: Cambridge University Press.

Hannah, S., Craig, J., Heckbert, J., Pentyliuk, M., Burke, K., & Siegel, L. (2009). *Reading Readiness Screening Tool*. Learning Disabilities Association of Alberta. Retrieved from www.righttoread.ca.

Hemphill, F. C., & Vanneman, A. (2011). *Achievement gaps: how Hispanic and white students in public schools perform in mathematics and reading on the national assessment of educational progress. Statistical analysis report*. NCES 2011-459. National Center for Education Statistics.

Hermanto, N., Moreno, S., & Bialystok, E. (2012). Linguistic and metalinguistic outcomes of intense immersion education: how bilingual?. *International journal of bilingual education and bilingualism*, *15*(2), 131-145.

Hirsh-Pasek, K., Gleitman, L. R., & Gleitman, H. (1978). What did the brain say to the mind? Study of the detection and report of ambiguity by young children. In A. Sinclair, R. J. Jarvella, & W. J. M. Levelt (Eds.), *The child's conception of language*. New York, NY: Springer Verlag.

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: implications for closing achievement gaps. *Developmental Psychology*, *49*(1), 4.

Hoover, W., & Gough, P. B. (1990). The Simple View of Reading. *Reading and Writing: An Interdisciplinary Journal*, *2*, 127–160.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L., & Trueswell, J. (2001). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research*, *29*, 597-626.

Johnson, C. J. (1992). Cognitive components of naming in children: Effects of referential uncertainty and stimulus realism. *Journal of Experimental Child Psychology*, *53*(1), 24-44.

Johnson, C. J., & Clark, J. M. (1988). Children's picture naming difficulty and errors: Effects of age of acquisition, uncertainty, and name generality. *Applied Psycholinguistics*, *9*(04), 351-365.

Joshi, R. M., & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology*, *21*(2), 85-97.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of educational Psychology*, *80*(4), 437.

Katz, N., Baker, E., & Macnamara, J. (1974). What's in a name? A study of how children learn common and proper names. *Child Development*, *45*, 469-473.

Keil, F. (1980). Development of the ability to perceive ambiguities: evidence for the task specificity of a linguistic skill. *Journal of Psycholinguistic Research*, *9*, 219-30.

Kendeou, P., Savage, R., & Van den Broek, P. (2009). Revisiting the simple view of reading. *British Journal of Educational Psychology*, *79*(2), 353-370.

Keppel, G., & Wickens, T. D., 1942. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Upper Saddle River, N.J: Pearson Prentice Hall.

Kidd, E., & Bavin, E. (2005). Lexical and referential cues to sentence interpretation: An investigation of children’s interpretations of ambiguous sentences. *Journal of Child Language*, *32*(4), 855-876.

Kidd, E., & Bavin, E. (2007). Lexical and referential influences on online spoken language comprehension: A comparison of adults and primary-school-age children. *First Language*, *27*(1), 29-52.

Kidd, E., Stewart, A., & Serratrice, L. (2011). Children do not overcome lexical biases where adults do: The role of the referential scene in garden-path recovery. *Journal of Child Language*, *38*, 222-234.

Liberman, I. Y., Shankweiler, D., Liberman, A. M., Fowler, C., & Fischer, F. W. (1974). Phonetic segmentation and recoding in the beginning reader. *Haskins Laboratories Status Report on Speech Research*, *39-40*, 105-121.

Lonigan, C. J., & Phillips, B. M. (2016). Response to instruction in preschool: Results of two randomized studies with children at significant risk of reading difficulties. *Journal of Educational Psychology*, *108*(1), 114.

McDaniel, D., & Cairns, H. S. (1990). The child as informant: Eliciting linguistic intuitions from young children. *Journal of Psycholinguistic Research, 19*, 331-334.

McDaniel, D., & Cairns, H. S. (1996). Eliciting judgments of grammaticality and reference. In D. McDaniel, C. McKee, & H.S. Cairns (Eds.) *Methods for Assessing Children's Syntax*, MIT Press.

Miller, J., & McKenna, M. C. (2016). *World literacy: How countries rank and why it matters*. New York: Routledge

Moore, R. S., & Moore, D. N. (1975). *Better late than early: a new approach to your child's education*. Reader's Digest Association.

Morais, J. (1991). Phonological awareness: a bridge between language and literacy. In D. Sawyer, & B. Fox (Eds.), *Phonological Awareness in Reading: The Evolution of Current Perspectives* (pp. 31–71). New York: Springer-Verlag.

Morais, J., Alegria, J., & Content, A. (1987). The relationships between segmental analysis and alphabetic literacy: an interactive view. *Cahiers de Psychologie Cognitive, 7*, 1–24.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7*, 323–331.

Moriguchi, Y., Okanda, M., & Itakura, S. (2008). Young children's yes bias: How does it relate to verbal ability, inhibitory control, and theory of mind?. *First language*, *28*(4), 431-442.

Morris, D., & Ebooks Corporation. (2014). *Morris informal reading inventory: Preprimer through grade 8*. New York, NY: The Guilford Press.

Nation, K., & Snowling, M. J. (2000). Factors influencing syntactic awareness skills in normal readers and poor comprehenders. *Applied Psycholinguistics*, *21*(2), 229-241.

National Reading Panel. (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development (US), National Institutes of Health.

National Center for Education Statistics. (1999). *The condition of education*. US Dept. of Education, Office of Educational Research and Improvement, National Center for Education Statistics. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education.

National Center for Education Statistics. (2011). *Achievement gaps: How Hispanic and white students in public schools perform in mathematics and reading on the national assessment of educational progress: Highlights*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education.

Omaki, A., & Lidz, J. (2015). Linking parser development to acquisition of syntactic knowledge. *Language Acquisition*, *22*(2), 158-192.

Peterson, C., Dowden, C., & Tobin, J. (1999). Interviewing preschoolers: Comparisons of yes/no and wh-questions. *Law and Human Behavior*, *23,* 539–555.

Raschke, V. R. (2013). Processes underlying syntactic control: Evaluating linguistically diverse children. (Unpublished Doctoral Dissertation). Loyola University Chicago. Chicago, IL.

Roth, F. P., Speece, D. L., Cooper, D. H., & De La Paz, S. (1996). Unresolved mysteries how do metalinguistic and narrative skills connect with early reading?. *The Journal of Special Education*, *30*(3), 257-277.

Saddy, D. (1992). Sensitivity to islands in an aphasic individual. In H. Goodluck, & M. Rochemont (Eds.), *Island constraints: Theory, acquisition and processing.* Dordrecht, NL: Springer Netherlands.

Schlisselberg, G. (1988). Development of selected conservation skills and the ability to judge sentential well-formedness in young children. (Unpublished Doctoral Dissertation). City University of New York, New York, NY.

Scholl, D. M., & Ryan, E. B. (1980). Development of metalinguistic performance in the early school years. *Language and Speech*, *23*(2), 199-211.

Shakibai, M.T. (2008) The efficacy of a training program to teach kindergarteners lexical ambiguity detection. (Unpublished Doctoral Dissertation). City University of New York, New York, NY.

Shipley, E. F., Smith, C. S., & Gleitman, L. R. (1969). A study in the acquisition of language: Free responses to commands. *Language*, *45*, 322-342.

Shultz, T., & Pilon, R. (1973). Development of the ability to detect linguistic ambiguity. *Child Development*, *44*, 728-733.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, *49*, 238-299.

Simmons, D. C., Coyne, M. D., Kwok, O. M., McDonagh, S., Harn, B. A., & Kame'enui, E. J. (2008). Indexing response to intervention a longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities*, *41*(2), 158-173.

Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language*, *58*(2), 574-608.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174-215.

Snow, C., Burns, M.S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Spivey-Knowlton, M.J., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, *55*, 227-267.

Stallman, A. C., & Pearson, P. D. (1990). Formal measures of early literacy. *Center for the Study of Reading Technical Report; no. 511*.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360-407.

Stanovich, K. E., Cunningham, A. E., & Feeman, D. J. (1984). Intelligence, cognitive skills, and early reading progress. *Reading Research Quarterly, 19*(3), 278-303.

Steffensen, M.S. (1978). Satisfying inquisitive adults: Some simple methods of answering yes-no questions. *Journal of Child Language*, *5*, 221–236.

Stockard, J. (2010). Promoting reading achievement and countering the "fourth-grade slump": The impact of direct instruction on reading achievement in fifth grade. *Journal of Education for Students Placed at Risk (JESPAR), 15*(3), 218-240.

Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children. *Reading and Writing*, *4*(3), 245-256.

Székely, A., & Bates, E. (2000). Objective visual complexity as a variable in studies of picture naming. *Center for Research in Language Newsletter*, *12*(2).

Thornton, R. (1996). Elicited production. In D. McDaniel, C. McKee, & H.S. Cairns (Eds.). *Methods for assessing children's syntax*, Cambridge, MA: MIT Press.

Thomas, J. (1988). The role played by metalinguistic awareness in second and third language learning. *Journal of Multilingual & Multicultural Development*, *9*(3), 235-246.

Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading*, *32*(4), 383-401.

Tong, X., Deacon, S. H., & Cain, K. (2014). Morphological and syntactic awareness in poor comprehenders another piece of the puzzle. *Journal of Learning Disabilities, 47*(1), 22-33.

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology, 40*(1), 7-26.

Traxler, M. J. (2012). *Introduction to psycholinguistics: Understanding language science*. Malden, MA: Wiley-Blackwell.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying online sentence processing in young children. *Cognition, 73*(2), 89-134.

Trueswell, J. C., & Tanenhaus, M. K. (2005). *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*. Cambridge, MA: MIT Press.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*(3), 285-318.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*(4), 1067-1084.

Tunmer, W. E., & Herriman, M. L. (1984). The development of metalinguistic awareness: A conceptual overview. In W. Tunmer, C. Pratt, & M Herriman (Eds.), *Metalinguistic awareness in children: Theory, research, and implications.* Berlin, Heidelberg: Springer Berlin Heidelberg.

Tunmer, W.E., Pratt, C., & Herriman, M.L. (1984). Metalinguistic abilities and beginning reading. *Reading Research Quarterly, 2*, 134-158.

van Kleeck, A. (1994). Metalinguistic development. In G. Wallach & K. Butler (Eds.). *Language learning disabilities in school-age children and adolescents: Some principles and applications.* New York: Macmillan.

Vanneman, A., Hamilton, L., Anderson, J. B., & Rahman, T. (2009). *Achievement gaps: how black and white students in public schools perform in mathematics and reading on the national assessment of educational progress. Statistical analysis report.* NCES 2009-455. National Center for Education Statistics.

Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1998). The case for early intervention in diagnosing specific reading disability. *Journal of School Psychology, 36*(4), 367-397.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*(2), 192.

Wankoff, L. S. (1983). Selected metalinguistic variables as they relate to conservation and reading achievement in normal achieving and some learning-disabled children. (Unpublished Doctoral Dissertation). City University of New York, New York, NY.

Wankoff, L. S., & Cairns, H. S. (2009.) Why ambiguity detection is a predictor of early reading skill?. *Communication Disorders Quarterly, 30*, 183-192.

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology, 102*(1), 43.

Wimmer, H., Landerl, K., Linortner, R., & Hummer, P. (1991). The relationship of phonemic awareness to reading acquisition: More consequence than precondition but still important. *Cognition*, *40*(3), 219-249.

Yuill, N. (2009). The relation between ambiguity understanding and metalinguistic discussion of joking riddles in good and poor comprehenders: Potential for intervention and possible processes of change. *First Language, 29*(1), 65-79.

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science, 21*(4), 551-559.

Zipke, M. (2011). First graders receive instruction in homonym detection and meaning articulation: The effect of explicit metalinguistic awareness practice on beginning readers. *Reading Psychology, 32*(4), 349-371.

Zipke, M., Ehri, L. C., & Cairns, H. S. (2009). Using semantic ambiguity instruction to improve third graders' metalinguistic awareness and reading comprehension: An experimental study. *Reading Research Quarterly, 44*, 300-321.

Appendix A. Children's Comprehension of Two Types of Syntactic Ambiguity (Article 1)

Abstract

This study asks whether children accept both interpretations of ambiguous sentences with contexts supporting each option. Twenty-six 3- to 5-year-old English-speaking children and a control group of 30 English-speaking adults participated in a truth value judgment task. As a step towards evaluating the complexity of syntactic ambiguity, the materials included prepositional phrase ambiguities (e.g., *The boy saw the girl with binoculars*.) and reduced relative clause ambiguities (e.g., *The boy saw the girl using a magnifying glass*.). Children generally accepted both interpretations of such ambiguities. This suggests that children this age have not only mastered the syntactic configurations associated with these ambiguities, but they can appropriately assign multiple possible structural analyses to an individual expression. Children did reject more interpretations than adults. This may be due to differences in the developing parser and the adult parser.

Keywords: language development, language acquisition, syntactic ambiguity, structural ambiguity, comprehension, parsing, prepositional phrase, reduced relative clause, truth value judgement

.

1.0 Introduction

The study of ambiguity is a productive and influential avenue of research among scholars of sentence processing and comprehension. This is, in part, because ambiguities slow down the parsing mechanism, allowing researchers to better distinguish the underlying processing operations (Spivey-Knowlton & Sedivy, 1995). Ambiguity introduces complexity in comprehension as well because it requires the listener to assign one of multiple possible structural analyses to an utterance. While there are several types of ambiguity in natural language, the focus of this study is structural, or syntactic, in which different possible structural configurations yield a sentence with correspondingly more than one meaning. This study probes 3- to 5-year-olds' comprehension of sentences with prepositional phrase (PP) attachment ambiguities and reduced relative clause (RC) attachment ambiguities, testing whether individual children entertain two interpretations of each ambiguity type.

Previous research suggests that children command the individual structures associated with these ambiguity types as early as 2-years-old: prepositional phrases (Johnston & Slobin, 1979; Valian, 1986; Tomasello, 1987) and relative clauses (Limber, 1973; Hamburger, 1980; McKee, McDaniel, & Snedeker, 1998). Several studies have probed children's ambiguity processing focusing on the informational cues that inform ambiguity resolution (Trueswell, Sekerina, Hill, & Logrip, 1999; Snedeker & Trueswell, 2004; Kidd & Bavin, 2007; Snedeker & Yuan, 2008; Kidd, Stewart, & Serratrice, 2011). However, no study has tested directly whether children will accept both interpretations of an ambiguous expression given the appropriate discourse contexts. It is important to find out whether individual learners accept more than one meaning for the same word string because this requires more or a different kind of processing than when that word string has only one grammatical configuration. In other words, associating an utterance with two different structures creates additional processing load. Because of this, it is possible that this ability develops well after a child has mastered the individual structures associated with an ambiguity. It is not a given that 3- to 5-year-olds have both interpretations of an ambiguous sentence available to them. In fact, there may be variation across individuals. If so, this would speak to issues that matter to both the literature on syntactic development and parsing development. If some children show a greater capacity for computing both interpretations of an ambiguity than others, it would suggest a developmental trend that could be influenced by a variety of factors, such as children's use of limited informational cues during parsing.

This study is the first that I know of to probe children's comprehension of two syntactic ambiguities. PP ambiguity was included in this study because it is the focus of the majority of research on children's ambiguity comprehension and processing. RC ambiguity was included because it has comparable attachment sites to PP ambiguities but could be considered more complex as it is bi-clausal. Thus, this study asks whether participants accept both interpretations of two types of ambiguity with similar attachment sites but different degrees of syntactic complexity. Investigating children's performance on more than one ambiguity type is important if our goal is to discover parsing and grammatical principles that apply beyond individual structures.

The paper begins with a brief description of the two ambiguity types and then motivates the study considering its implications for research on syntactic and parsing development.

1.1 Ambiguity types

PP ambiguities and RC ambiguities have similar attachment sites. These occur when a prepositional phrase or a reduced relative clause can attach either to the local noun phrase (NP) as a modifier or to the verb phrase (VP) as an instrument. PP ambiguity is shown in (1) and (2). RC ambiguity is shown in (3) and (4).

(1) The boy [saw [the girl with binoculars]].
(2) The boy [saw [the girl] with binoculars].
(3) The boy [saw [the girl [using a magnifying glass]]].
(4) The boy [saw [the girl] [using a magnifying glass]].

Consider the PP ambiguity in (1) and (2). When the PP *with binoculars* is attached to the local NP *the girl* as in (1), the interpretation of the sentence is that the boy saw a girl who has binoculars. This interpretation will be referred to as the NP-attachment interpretation. When the PP is attached to the VP as in (2), the interpretation is that the boy used binoculars to see the girl. This interpretation will be referred to as the VP-attachment interpretation. Figure 1 illustrates the two interpretations, (1) on the right side and (2) on the left. Now consider the RC ambiguity in (3) and (4). In this case, when the reduced relative clause *using a magnifying glass* modifies the local NP *the girl* as in (3), the interpretation is that the boy saw a girl and that girl is using a magnifying glass. When the reduced relative clause attaches to the VP as in (4), the interpretation is that the boy used a magnifying glass to see the girl. Figure 2 illustrates its two interpretations, (3) on the right side and (4) on the left.

Figure 1: *The boy saw the girl with binoculars.*

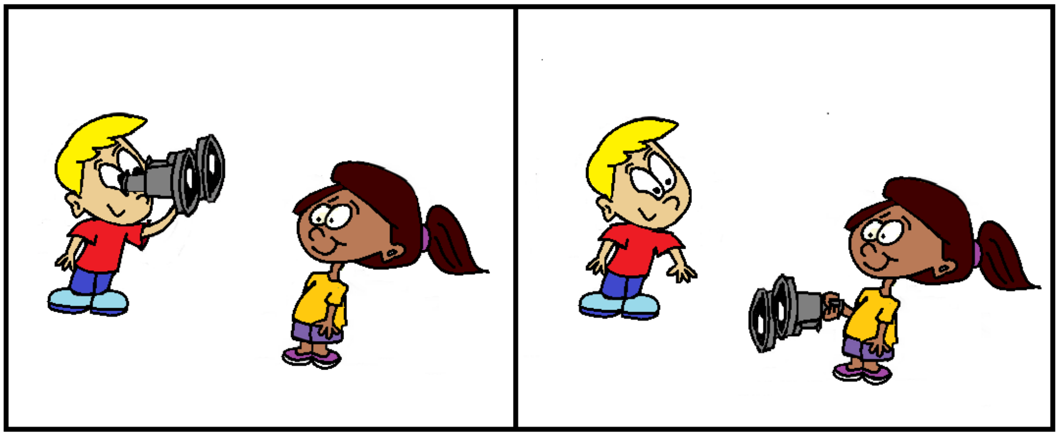


Figure 2. *The boy saw the girl using a magnifying glass.*

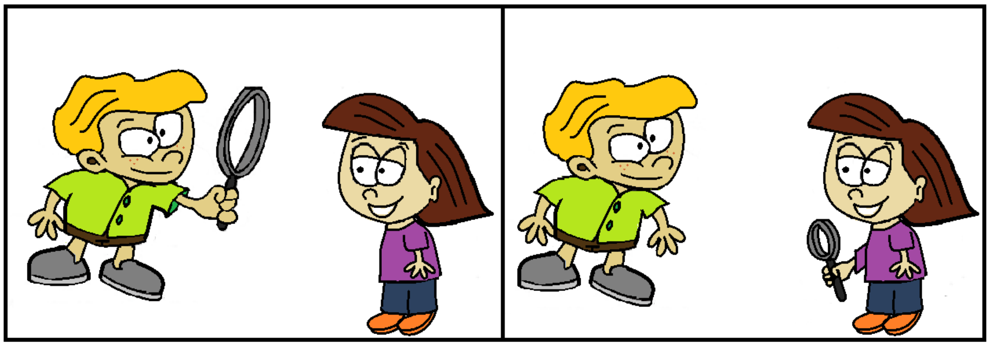


Table 1 summarizes the interpretation paradigm for each ambiguity type. The final column shows which example sentences above correspond to each interpretation.

Table 1: Summary of ambiguity types.

Ambiguity Type	**Attachment**	**Interpretation**	**Example Sentences**
Prepositional Phrase (PP)	NP-attachment	Modifier	(1)
	VP-attachment	Instrumental	(2)
Reduced Relative Clause (RC)	NP-attachment	Modifier	(3)
	VP-attachment	Instrumental	(4)

1.2 Syntactic development

The literature on children's comprehension of syntactic ambiguities assumes that all interpretations of a syntactic ambiguity are available to each child, but this has never been directly tested. This assumption seems reasonable because children's syntactic development occurs rapidly and at an early age. By the age of five, children's syntactic production and comprehension are in many ways almost adult-like (Crain & Fodor, 1989). Given that comprehension usually precedes production, examining the development of PPs and relative clauses in the production literature will serve to illustrate why the participants in the present study (aged 3;6-5;10) should be able to comprehend both relevant structure types.

Beginning with prepositional phrases, several studies have shown that children produce PPs as early as 2-years-old. For example, Tomasello (1987) reported a diary study of a child who produced prepositions as holophrases as early as 1;5 and was producing locative PPs before her second birthday. Valian (1986) looked at naturalistic data as well as some elicited production data to argue that 2- to 3-year-olds have certain syntactic categories. She found that 2-year-olds produce prepositional phrases in a variety of syntactic contexts (e.g., after verbs, after direct objects, etc.) and make relatively few errors in the process. Because of this, Valian argued that 2-year-olds have a conceptual representation of the category PP. Johnston and Slobin (1979) used an elicited production task to show that 2- to 4-year-old speakers of four different languages can produce locative expressions in their respective languages. English speakers in their study produced certain PPs as early as 2-years-old (e.g., *in, on*). Johnston and Slobin argued that there is a cross-linguistic order of acquisition of locative relations. English-speaking participants in their study were producing all the PPs in their proposed order before the age of five.

Similarly, several studies have shown that children produce relative clauses as young as 2-years-old. In a diary study of a single 2-year-old, Hamburger (1980) found what he called precursors to relative clauses. Limber (1973) collected spontaneous production and elicited production data from 11 children. He found that children were using relative clauses as young as 2;9. Similarly, in an elicited production study, McKee et al. (1998) found that 2-year-olds produce mostly adult-like relative clauses. Their errors did not point to syntactic incompetence, but were better analyzed as lexical and morphological errors. Importantly, their results showed that very young children can reliably produce reduced relative clauses of the type included in the current study. Because comprehension generally precedes production, these findings suggest that children in the current study should be able to comprehend both relevant structure types.

However, for a child to accept both interpretations of an ambiguous utterance, he or she must employ more than basic grammatical knowledge of a structure. He or she must be able to attach two different structural analyses to one string of words. If a child cannot do this, one could argue that he or she has not mastered the relevant syntax of an ambiguous expression. In fact, the complexity of ambiguity may disrupt the typical disparity between production and comprehension. As a reviewer pointed out, comprehension may not precede production in the case of ambiguity. Speakers knows their intended message and, thus, should be unaffected when producing ambiguity. Listeners, however, must reconstruct the intended meaning from the ambiguous signal, causing processing difficulty. For example, Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell (2001), using a combined act-out task and elicited production task, found that 4- to 5-year-olds could produce restrictive prepositional phrase modifiers to select a member of a set but had difficulty comprehending such modifiers. This supports the notion that the comprehension of ambiguity requires more than mastery of an individual grammatical structure. In spite of this complexity, the literature on children's ambiguity processing hints that 4- to 5-year-olds can access both interpretations of an ambiguity (Trueswell et al., 1999; Snedeker & Trueswell, 2004; Snedeker & Yuan, 2008; Kidd et al., 2011). For example, in eye-tracking studies such as Trueswell et al. (1999) and Snedeker and Trueswell (2004), children looked at varying lengths at objects that represented both interpretations of ambiguous sentences. However, this research was primarily concerned with the informational cues that influence children's ambiguity resolution. It did not explicitly test whether individual children could access both interpretations.

The first hypothesis of the current study is that young children often do access both interpretations of a syntactically ambiguous sentence, though this access is not necessarily conscious. Thus, in a comprehension task, they will accept both meanings of a structural ambiguity in most cases. However, the literature on both adults' and children's processing of syntactic ambiguities has shown that ambiguity resolution occurs so rapidly during processing that hearers will often not notice that an utterance was ambiguous (Frazier, 1978; Altmann, 1986). This rapid resolution informs a second hypothesis that sometimes both adults and children will only compute one interpretation and thus reject a true interpretation of an ambiguous sentence even when it is supported by the discourse context.

1.3 The developing parser

While the current study does not test parsing per se, it speaks to the development of parsing. For most theories, structural ambiguity creates additional processing load. Parallel models predict that the parser must compute all possible structures of an ambiguous sentence simultaneously, creating a load on the working memory (Macdonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994). Serial models predict that the parser must restart if it first computes the wrong structure (Frazier, 1978). Either way, the parser must do some extra work when it encounters word sequences like in (1) and (3).

Testing children's response to such a processing load has revealed certain differences between the developing parser and the adult parser. These differences suggest that children and adults may perform differently in an ambiguity comprehension task. Specifically, there are two differences that speak to the current study. First, even though the incremental nature of parsing is the same for adults and children (Borovsky, Elman, & Fernald, 2012; Mani & Huettig, 2012), children are less likely to revise an initial parse than adults are even when another parse is supported by many extra-linguistic cues (Trueswell et al., 1999; Snedeker & Trueswell, 2004; Kidd et al., 2011; Omaki & Lidz, 2015). This suggests that children in the current study should be less likely to accept both interpretations of an ambiguous sentence than adults. Specifically, if their initial parse does not correspond to the one supported by the discourse context, they should be less likely to reconsider their first interpretation when presented with conflicting context cues. To exemplify, Trueswell et al. (1999) paired an act-out task with eye-tracking to test children's processing of temporarily ambiguous sentences such as in (5).

(5) Put the frog on the napkin in the box.

Children's errors in the act-out task revealed that they did not revise their initial syntactic commitments (e.g., putting the distractor animal on the napkin, hopping the distractor animal to the napkin first and then to the box). This suggests that, once children parsed the first PP as a destination, they could not revise their parse even after hearing the second PP. Kidd et al. (2011) further supported this using a similar task with globally ambiguous sentences.

Another key difference between the adult parser and the developing parser is children's use of fewer informational cues during processing. Adults use a variety of cues when parsing ambiguous sentence. These include referential information (Altmann & Steedman, 1988), plausibility (Chambers, Tanenhaus, & Magnesun, 2004), lexical biases (Spivey-Knowlton & Sedivy, 1995), and more. However, experimental data suggests that children rely primarily on lexical co-occurrence biases when resolving PP ambiguities (Trueswell et al., 1999; Snedeker & Trueswell, 2004; Kidd & Bavin, 2005). These biases cause adults and children to have preferences for one interpretation of an ambiguous construction over another based on the semantic category of the verb in the sentence (Spivey-Knowlton & Sedivy, 1995; Snedeker & Trueswell, 2004; Kidd & Bavin, 2005). For example, verbs expressing an action, such as *hit* or *poke*, more frequently co-occur with VP-attachment, whereas perception or so-called psych verbs, such as *listen* or *see*, more frequently co-occur with NP-attachment. Because of this, people tend to parse a PP in a sentence with an action verb as having VP-attachment and a PP in a sentence with a perception verb as having NP-attachment. Several studies suggest that children's parsing relies primarily on these biases.

For example, Snedeker and Trueswell (2004) paired eye-tracking with an act-out task to test 5-year-olds' sensitivity to verb biases. They found that children's eye gaze patterns were influenced by verb bias but not referential scene when the verbs were biased towards VP-attachment or NP-attachment. Only in the condition where the verbs were considered equi-biased did they find a weak effect of referential context. This suggests that children may attend to weaker constraints in the absence of strong verb biases. Kidd et al. (2011) also paired an act-out task with eye-tracking to test whether instrument plausibility could over-ride verb biases in 5-year-olds' processing of sentences like (6).

(6) Cut the cake with the candle.

The verb *cut* is biased towards VP-attachment (instrumental interpretation). However, a candle is an implausible instrument for cutting a cake but a very plausible modifier. In an act-out task, children responded based on verb bias and generally ignored plausibility (i.e., they used the candle to cut the cake).

Overall, research on children's parsing suggests that they apply fewer informational cues than adults do while processing ambiguities. This supports the prediction that children will reject more possible interpretations than adults will. Adults, who incorporate many cues into their parse, will be able to use extra-linguistic cues to guide them to the parse supported by context of the story. Because of this, they should be more likely to accept both interpretations of an ambiguous sentence when presented with discourse contexts that support each interpretation. Children who rely mostly on lexical cues should come to the same parse regardless of the discourse context. If they don't reanalyze the sentence at a later point, this could cause them to reject true possible interpretations of a sentence.

However, it should be noted that the body of literature on children's parsing of syntactic ambiguity is still small. There are several informational cues that have not yet been probed. For example, several studies have probed the effect of verb biases in children's parsing, but to my knowledge no study has probed other types of lexical biases yet. For example, Trueswell and Tanenhaus (2005) point out that prepositions themselves have biases towards NP-attachment or VP-attachment. No study has looked at the effect of preposition bias in children. There is also very little research on children's processing of different types of syntactic ambiguity. Felser, Marinis, and Clahsen (2003) studied the effect of working memory on children's processing of a type of relative clause attachment ambiguity. Aside from this study, the entire literature on children's syntactic ambiguity parsing has focused on prepositional phrase attachment ambiguities. Because of this dearth of research, we do not know yet whether children use more informational cues than research has so far identified.

In sum, these differences between the adult parser and developing parser suggest that children and adults may perform differently in the present study with respect to how frequently they reject or accept certain interpretations of syntactically ambiguous sentences. Because children rely heavily on fewer cues and are less able to revise an initial parse, this suggests that children may reject more possible interpretations of ambiguous sentences than adults. The following section describes the current study's methods for testing whether children access both interpretations of a syntactic ambiguity.

2.0 Methods

To make the methods clearer in the subsequent sections, this section begins with an introduction to the experimental task and conditions before describing the rest of the experimental design. This study used a truth value judgment (TVJ) task (Abrams, Chiarello, Cress, Green, & Ellelt, 1978; Crain & McKee, 1985; Gordon & Chafetz, 1986). This task asks participants to judge whether a statement is true according to some context that was previously described in a story or preamble (Gordon, 1996). In other words, participants judge whether or not a meaning-utterance pair is true. One important property of this task is that it can test whether participants will accept the same meaning with multiple utterances or the same utterance with multiple meanings. Therefore, this task is well suited to the current study's research question, which asks whether children will accept two meanings for the same utterance.

In the current study, participants were asked to make judgments about whether an ambiguous utterance was true or not in three different interpretation contexts, two in which each of the possible interpretations were true and one in which the interpretation was false. Each of these meaning/utterance pairs makes up an experimental item. In other words, each experimental item includes both the utterance and the interpretation that is supported by the story associated with the utterance. The same utterance occurs in three different items each testing one of the interpretation conditions. These utterances correspond to the two ambiguity types introduced in the Section 1.1 - PP and RC.

2.1 Participants

The participants were 26 children aged 3;6- 5;10 (mean age 4;10), and a control group of 30 adults. The children were tested at two university-area preschools in Tucson, Arizona. All participants were native speakers of English; all but two children were monolingual. The adult group consisted of undergraduates at the University of Arizona.

2.2 Materials

This experiment has three types of item: teaching, experimental, and filler. Four teaching utterances were associated with a teaching story at the beginning of the experiment. The experimenters gave the children some feedback on these items if they appeared to not understand the task. Responses to these items were not analyzed or treated as a pre-test.

Each experimental item had two parts: an experimental utterance and a supported interpretation. As shown in Table 2, three experimental utterances represented each of the two ambiguity types, totaling to six experimental utterances.

Table 2. Experimental utterances by ambiguity type.[1]

Experimental Utterances by Ambiguity Type	
Prepositional Phrase (PP)	PP1: The man poked the monkey with a banana. PP2: The man pointed to the girl with an umbrella. PP3: The boy saw the girl with binoculars.
Reduced Relative Clause (RC)	RC1: The boy poked the girl using a pencil. RC2: The lady pointed to the man using a baseball bat. RC3: The boy saw the girl using a magnifying glass.

Each ambiguity type included three utterances. The same three verbs were used for both ambiguity types: *poke, point to,* and *see*. As mentioned in the literature review, verbs can be biased towards VP-attachment or NP-attachment based on their lexical co-occurrence frequencies. Action verbs are thought to be biased towards VP-attachment; perception verbs are thought to be biased towards NP-attachment (Spivey-Knowlton & Sedivy, 1995; Snedeker & Trueswell, 2004). Thus, the verbs *poke* and *point to* should be biased towards VP-attachment, and *see* should be biased towards NP-attachment.

Each experimental utterance was paired with three different supporting interpretations, and therefore was repeated after three different stories. Two stories supported each of the possible true interpretations (i.e., one scenario supported NP-attachment while another supported VP-attachment). In the third story, the utterance was false regardless of interpretation. This totaled to 18 items (6 utterances x 3 interpretations).[2] Two different experimental items (one

---

[1] All experimental utterances were pilot tested with six children as young as 3 to ensure that all lexical items were familiar to children of the age group in this study. Utterances were not controlled for frequency or length.

[2] The original experiment included a third ambiguity type that occurs when an adjective precedes a conjoined NP (e.g., *The teacher found green frogs and fish*). The adjective could either modify the entire conjoined NP or only the first NP. There were three of these items, making the actual total number of experimental items 27. Data from these adjective items were excluded here because of questions regarding the interpretation of the participants' responses. Because the first NP is modified in both interpretations, a participant could theoretically respond "yes" in both true conditions even if she had only computed the interpretation where the first NP is modified.

from each ambiguity type) were tested in each story. Thus, participants heard nine experimental stories and one teaching story for a total of ten stories. The stories and items were heard in one fixed order. To address ordering and priming effects, the interpretation that was supported for each ambiguity type was rotated each story. This made it so that any particular ambiguity type never had the same interpretation two stories in a row. Additionally, a PP ambiguity and a RC ambiguity never had the same interpretation in the same story.

In addition to the experimental items, there were four filler items per story. The filler items were simple utterances of varying structures. They provided an evaluation of whether children were attending to the task or had any sort of strong response bias. The fillers were designed to balance out the number of “yes” and “no” responses that the child was expected to give. Throughout the experiment, participants heard a total of 58 items (18 experimental + 36 filler + 4 teaching = 58). For a list of all experimental and filler items, see Appendix A.

2.3 Procedure

The experiment lasted between 20 and 30 minutes, depending on attention of the child. Participants were presented the story as a game. The first experimenter told the contextual stories with a set of toys. After each of the nine experimental stories, a puppet controlled by a second experimenter used six statements to describe the story. These statements included the two ambiguous experimental utterances, which were buried in four filler items. Participants “helped” the puppet by telling him whether he had said the right thing or wrong thing based on what had occurred in the story. All of the puppet’s statements were grammatical to ensure that judgments were based on interpretation only and not linguistic form.[3]

Stories were constructed so that there was only one reasonable interpretation of each ambiguous experimental utterance in each story. To test whether participants would accept both interpretations of an ambiguous sentence, every experimental utterance was presented after two different stories, each supporting one of the different possible interpretations of the sentence. Every utterance was also presented after a scenario where it would be considered false based on either interpretation. Take, for example, the following experimental utterance, repeated from (1).

(7) The boy saw the girl with binoculars.

Participants heard it after a story involving a boy who was using binoculars to see a girl (i.e., VP-attachment). At another point, participants heard the same utterance after a story involving a girl with binoculars (i.e., NP-attachment). Participants who responded “yes” to the statement after both stories demonstrated that they accepted the meaning-utterance pair as true in both scenarios, thereby indicating that they accessed both interpretations of the ambiguity. If participants

---

[3] Experimenters spent several hours practicing the delivery of the stories and the experimental utterances. This was to ensure that the stories were consistent across children and that the experimental utterances were delivered with neutral prosody.

accepted the meaning-utterance pair after one story but not the other, this suggests that they accessed only one of the possible meanings of the utterance in question. For an example of a contextual story, see Appendix B.

Experimental items were also included in scenarios where they could only be interpreted as false. For example, the utterance in (7) was heard after a scenario in which a boy is looking at a girl but a man has the binoculars rather than a girl. Including such items was a control to ensure that children are capable of processing the falsehood of these types of sentences. This was important because young children's "yes" bias can compromise an experiment like this one. Children respond "yes" more frequently than "no" during an experiment because they feel uncomfortable telling an unfamiliar adult that he or she is wrong (Steffensen, 1978; Peterson, Dowden, & Tobin, 1999). When children respond "no" to an item in the scenario in which it is truly false, this gives us confidence that their positive responses in the other scenarios are a valid reflection of their sentence comprehension and not simply a preference for responding "yes". The scenarios were also constructed so that both "yes" and "no" responses would be pragmatically felicitous. For example, plausible deniability was considered (Crain et al., 1996).

2.4 Data treatment

Responses were recorded by hand on individual session sheets by the experimenter who was also controlling the puppet. This data was then coded based on percentage of "yes" responses as well as percentage of expected response. Expected response differed by item, depending on the supporting context of each story. For the experimental items, the expected response was "yes" in both cases where the story supported one of the possible ambiguous interpretations and "no" in the false condition. For filler items, the expected response varied by story. Every story either had two true and two false fillers or one true and three false fillers.

3.0 Results

All participants who completed the study are included in the analyses below. I will begin with performance on the filler items. All but three children were correct on at least 90% of the fillers. The lowest percent correct for a child on the fillers was 83%. Since 25 of the 36 fillers had a target "no" response, this suggests that no children had a significant "yes" bias. It also shows that they understood the task and were generally attending to the materials throughout the game. Adults performed better on the filler items than the children. The lowest percent correct for an adult on the fillers was 91%.

Performance on the experimental items is detailed in Table 3. First, adults and children performed similarly on the NP-attachment interpretation of the RC items. But adults accepted both interpretations of the PP utterances more than children. Adults also accepted the VP-attachment interpretations of the RC utterances more than children. These differences between adults and children will be further discussed in the Section 4.0.

Table 3. Percentage of "yes" responses by ambiguous type and supported interpretation.

Ambiguity Type	**Interpretation**	**Adults**	**Children**
PP	NP-Attachment	72.2%	59.0%
	VP-Attachment	93.3%	83.3%
	False	11.1%	24.4%
RC	NP-Attachment	60.0%	62.8%
	VP-Attachment	95.6%	78.2%
	False	0%	7.7%

Table 3 gives the raw percentages, which suggest that adults and children frequently say "yes" to both the NP-attachment and VP-attachment interpretations. However, as previously mentioned, yes bias can be problematic in this type of experiment. To account for any such bias, A' values were calculated. The A' equation takes into account the hit rate ("yes" responses to true interpretations) and the false alarm rate ("yes" responses to false items) to produce a value between 0 and 1 (Zhang & Mueller, 2005). An A' of 1 means that the participant perfectly discriminated true items from false items. In other words, he or she said "yes" to every true item and "no" to every false item. An A' value of 0.5 mean the participant is performing at chance. Table 4 shows the average A' value for children and adults separately.

Table 4. Mean A' value for children and adults divided by ambiguity type.

	Adult	**Children**
PP Ambiguity	0.90	0.80
RC Ambiguity	0.94	0.92
Total	0.90	0.85

The mean A' values in Table 4 show that both children and adults discriminated the true items from false items well, although children had lower mean A' scores than adults on the PP items.

It is also important to address certain differences between items. For both ambiguity types, adults and children generally responded "yes" to both interpretations for the two items containing *see* and *point to.* However, for each ambiguity type, the utterance containing the verb *poke* behaved differently than the other two. Very few children and adults accepted both interpretations of the *poke* utterances. Rather, the majority of participants rejected the NP-attachment interpretation and accepted the VP-attachment interpretation. This suggests that both groups had a strong VP-attachment preference on these items, as shown in Table 4.[4]

[4] In Table 4, the percentages add to greater than 100% on three rows because one child answered "no" to both interpretations on PP3 and RC3 and one adult answered "no" to both interpretations on RC 3.

Table 5. Percentage of responses on individual items.

PP Items		**Yes to both interpretations**	**No to VP-Attachment**	**No to NP-attachment**
PP1: The man poked the monkey with a banana.	Children	15.4%	0%	84.6%
	Adults	20%	0%	80%
PP2: The man pointed to the girl with an umbrella.	Children	53.9%	19.2%	26.9%
	Adults	93.3%	3.3%	3.3%
PP3: The boy saw the girl with binoculars.	Children	61.5%	30.8%	11.5%
	Adults	83.3%	16.7%	0%
RC Items		**Yes to both interpretations**	**No to VP-Attachment**	**No to NP-attachment**
RC1: The boy poked the girl using a pencil.	Children	30.8%	3.9%	65.4%
	Adults	33.3%	3.3%	66.7%
RC2: The lady pointed to the man using a baseball bat.	Children	46.1%	38.5%	15.4%
	Adults	76.67%	6.67%	16.67%
RC3: The boy saw the girl using a magnifying glass.	Children	50%	23.1%	30.8%
	Adults	63.3%	3.3%	36.7%

These differences between items are likely due to the previously mentioned verb biases. The verb *poke* expresses an action and is thus biased towards VP-attachment, which is likely why participants showed such a strong preference for the VP-attachment interpretation. While the literature indicates that *see* has an NP-attachment bias and *point to* has a VP-attachment bias, these biases may not have been robust enough to strongly influence responses in a relatively slow off-line task. Interestingly, even though the research on verb biases has emphasized PP ambiguities, we see comparable patterns between PP items and RC items that contain the same verb. This could indicate that verb biases influence syntactic attachment in general and are not limited to a particular construction. However, it is important to note that both the preposition *with* and the gerund *using* may be biased in themselves towards an instrument interpretation, as suggested by Trueswell and Tanenhaus (2005). This could also account for the comparable patterns in the two syntactic structures.

The results also show individual variation among children. We can see this in the number of experimental utterances for which each a child accepted both interpretations. On the low end, three of the 26 children accepted both meanings for none of the six experimental utterance. Two of these children were 5-years-old; one was 4-years-old. On the high end, three children accepted both meanings for five of the six utterances. Two of these children were 4-years-old; one was 3-years-old. This suggests that the older children were not more willing to accept both interpretations of an utterance than the younger children. This is supported by the analyses in the following section.

4.0 Analyses

The data was analyzed using generalized linear mixed effects logistic models (GLMEs). All analyses were conducted in R (R Core Team, 2014) using the lme4 package (Bates, Maechler, Bolker, & Walker, 2014). All models used participants' yes/no responses as the dependent measure. GLMEs were chosen because of their ability to handle binomial distributions for a binary dependent measure. For all models, subject and item were included as random intercepts. Item was included as a random intercept because the order of items was the same for all participants. Thus, the model can account for item ordering effects. There were no random slopes in the models.

The first model compares children's and adults' responses on items in false contexts to their responses on items in both true interpretation contexts together (NP-attachment and VP-attachment). The percentage of "yes" responses to items in false contexts represents participants' bias to respond "yes" to a statement in the experimental setting. I predict that participants will respond "yes" to the true interpretations significantly more than to the false interpretation, thereby indicating that they are accessing both interpretations. The model included age group (child or adults), interpretation context (true or false), and the interaction as fixed effects. The interaction was significant (Est.=-2.12, SE =0.47, z=-4.57, p<.001). Both simple effects of interpretation context were significant. Children responded "yes" to the true contexts more than to the false context (Est.=3.42, SE=0.80, z=4.25, p<.001), and so did the adults (Est.=5.54, SE=0.85, z=6.50, p<.001). Both simple effects of age group were also significant. Children responded "yes" in true contexts less than the adults (Est.=-0.77, SE=0.25, z=-3.08, p<.01), but children responded "yes" in false contexts more than adults (Est.=1.35, SE=0.42, z=3.18, p<.01).

The second model compared children's and adults' responses to only the two true interpretation contexts. Thus, this model tested whether adults and children responded "yes" more frequently to items in an NP-attachment context or a VP-attachment context. The model included age group (child or adults), interpretation context (VP-attachment and NP-attachment), and the interaction as fixed effects. Again, subject and item were random intercepts. The interaction was again significant (Est.=-1.21, SE=0.48, z=-2.52, p<.05). The simple effect of interpretation context was significant only for adults. They said "yes" in VP-interpretation contexts more than in NP-interpretation contexts (Est.=2.47, SE=0.91, z=2.71, p<.01). Children trended in that same direction, but it was not significant (Est.=1.26, SE=0.87, z=1.44, p>.05). The simple effect of group was only significant for VP-attachment contexts. Adults responded "yes" to VP-attachment interpretations more than children (Est.=1.57, SE=0.42, z=3.74, p<.001). There was no significant difference between groups for NP-attachment contexts (Est.=-0.36, SE=0.32, z=-0.13, p>.05).

The final model only included children to allow for an analysis of the effect of children's age in months on their responses. It compares children's responses on items in false contexts to their responses on items in both true interpretation contexts together (NP-attachment and VP-attachment). Thus, this model investigates whether children say "yes" to true interpretations more often as they get older and say "yes" to false interpretations less often as they get older.

The model included age in months as a continuous variable, interpretation context (true or false), and the interaction as fixed effects. Again, subject and item were random intercepts. The interaction was not significant (Est.=-0.04, SE=0.04, z= -1.0, p>.05). The simple effect of age in true contexts was significant. Interestingly, older children responded “yes” to the true interpretations less than the younger children (Est.=-0.05, SE= 0.02, z=-2.27, p< 0.05). Older children also responded “yes” to the false items less than younger children, but this trend was not significant (Est.=-0.01, SE= 0.03, z= -0.34, p>.05).

These analyses reveal several notable findings. First, children and adults responded “yes” to both true interpretations significantly more than they did to the false interpretation. This suggests that children and adults generally accessed both interpretations of the ambiguity. Analyses did reveal some differences between children and adults. Adults said “yes” to true items more than children did, but adults said “yes” to false items less than children did. Furthermore, adults responded “yes” significantly more than children did to the VP-attachment interpretation items but not to the NP-attachment interpretation items. Finally, age affected children’s responses. Interestingly, older children were less likely than younger children to accept both true interpretations. In this respect, the younger children performed more like adults than the older children did. It is important to note here that younger children did not respond “yes” to the false items significantly more than older children. There was also no significant difference between younger children and older children on the filler items. In fact, the 3- to 4-year-olds had a slightly higher mean percent correct on filler items (97%) than the 5-year-olds (96%). Thus, the fact that younger children accepted both interpretations of a syntactic ambiguity is not likely due to a stronger yes bias in them than in the older children. If it were the case that the younger children’s yes bias was stronger than the older children’s, they would have responded “yes” more than older children to false items and filler items with a target “no” response. Rather, the best explanation for these results is that younger children were more disposed to accept both interpretations of the ambiguous items. The following section discusses the implications of these findings.

5.0 General discussion

This study shows that 3- to 5 year-olds can entertain both interpretations of a syntactically ambiguous sentence. They responded “yes” to both the true interpretations significantly more than to the false interpretation. Returning to the complexity point raised earlier, this is an important finding. It shows another way in which the parsers of 3- to 5-year-old children are sophisticated. In other words, these young parsers cannot only apply their syntactic knowledge in real-time interaction with lexical information, but they can also handle the additional processing load associated with structural ambiguity. Furthermore, adults and children generally accepted both interpretations for both PP and RC constructions. This suggests that the added complexity of the bi-clausal RC structure did not significantly affect participants’ ability to access both interpretations.

Although this study did not explicitly test parsing, it has implications for the study of children's parsing. The results showed that adults accepted both interpretations of an utterance more often than children. This may be due to differences in the adult parser and the developing parser. Previous research has shown that adults are more likely than children to attend to plausibility and referential cues (Trueswell et al., 1999; Snedeker & Trueswell, 2004; Kidd & Bavin, 2005) and they are more likely than children to revise a parse (Trueswell et al., 1999; Snedeker & Trueswell, 2004; Kidd et al., 2011). Either or both of these differences could have contributed to children rejecting more true interpretations than adults.

This begs the question of what is occurring when children accepted both interpretations of an ambiguous utterance in the current study. The previously mentioned differences between adult and child processing suggest two possibilities. The first is that children, because of their reliance on verb bias, initially parsed an ambiguous utterance the same in both scenarios but at some point revised their parse to suit the context. Because this task has no online data, the revision could have occurred a relatively long time after the utterance was first heard. It may be that children are less able to immediately revise a parse than adults, but they can revise parses in a slower task. The second possibility is that children parsed the sentence in two different ways depending on the context of the scenario, so no revision was needed. This would suggest that children used cues beyond verb bias to parse the sentence. The exception to this would be the two utterances containing the verb *poke*, which both elicited robust VP-attachment preferences. It may be that the verbs in the other four utterances, *see* and *point to*, have weaker biases than *poke*, causing children to draw on extra-linguistic context cues. This is supported by Snedeker and Trueswell (2004) who found that children make some use of referential cues in the absence of strong verb biases. Further online research is needed to gain a clearer picture of children's processing as they access both interpretations of a syntactic ambiguity.

These results also showed an interesting effect of age. The younger children in the study accepted both interpretations more often than the older children. The younger children were not significantly more likely than older children to say "yes" to filler items or false interpretation items, however. This suggests that the younger children were in fact responding differently to the ambiguity than older children. In the absence of online data, one can only speculate why this might be. A possible explanation relates again to children's verb biases. Kidd and Bavin (2005) used an act-out task with 3- and 5-year-olds to test the effect of verb bias on their comprehension of PP ambiguities. They found that 3-year-olds seemed to have a general preference for VP-attachment. In fact, they acted out the VP-attachment interpretation almost half the time in the condition where the verb was biased towards NP-attachment. Kidd and Bavin suggested that one possible explanation is that 3-year-olds are still learning the lexical co-occurrence restrictions associated with verbs that have an NP-attachment bias. This suggests a developmental trend in which older children have more established lexical biases than younger children. If it is true that older children have stronger verb biases than younger children, it follows that they may be more likely to reject a true interpretation that does not match their bias for that verb. Further research on the ambiguity processing of children 3-years-old and younger would elucidate this issue.

5.1 Future research

These findings have implications for research on children's metalinguistic awareness of syntactic ambiguity, an ability that links to reading development in 8- to 9-year-olds (Yuill, 2009). While younger children can detect both lexical and communicative ambiguity (Cairns, Waltzman, & Schlisselberg, 2004; Gillis & Nilsen, 2014), several scholars have concluded that children under 8-years-old cannot detect syntactic ambiguity (Wankoff, 1983; Cairns et al., 2004). It is unclear from these studies whether younger children's poor performance reflects a true lack of metalinguistic awareness or children's failure to compute both interpretations in the first place. The current study found that 3- to 5-year-olds usually accepted both interpretations of ambiguous sentences. Based on this, it is unlikely that children's apparent lack of metalinguistic awareness in previous research is due to a failure to comprehend both interpretations of an ambiguity. However, the data presented here did show a certain amount of variability, suggesting that some children may be more likely to compute both interpretations of a sentence than others. One might predict that children who consistently compute both interpretations in a comprehension task are likely to develop metalinguistic awareness of ambiguity earlier than those who do not. The current study also found children more likely to reject a possible interpretation of a sentence when it contained a verb with a strong lexical bias. This suggests that verb bias should be attended to in future research on children's detection of syntactic ambiguity.

Appendix A1: List of items by story

F=Fillers

Teaching Items:

1. The lady ate a strawberry.
2. The girl was riding in a wagon.
3. The girl watched the elephant eat a strawberry.
4. The lady watched the elephant eat a strawberry.

Story 1: Picnic

5. F: All these toys were at a picnic.
6. **RC:** The lady pointed to the man using a baseball bat.
7. F: This monkey brought a banana to the picnic.
8. F: The lady brought a baseball bat to the picnic.
9. **PP:** The man poked the monkey with a banana.
10. F: There were 8 pieces of cake.

Story 2: Science

11. F: All these toys were on a science trip.
12. **PP:** The boy saw the girl with binoculars.
13. F: This girl brought a pencil to the river.
14. F: There were 7 bugs.
15. **RC:** The boy poked the girl using a pencil.
16. F: The fish were green.

.

Story 3: Farm

17. F: All these toys went to a farm.
18. **PP:** The man pointed to the girl with an umbrella.
19. F: The ducks were dirty.
20. **RC:** The boy saw the girl using a magnifying glass.
21. F: There were six pigs.
22. F: The man brought an umbrella to the farm.

Story 4: Birthday Party

23. F: All these toys were at a birthday party.
24. **RC:** The lady pointed to the man using a baseball bat.
25. F: There were 5 monkeys at the party.
26. F: This man brought a baseball bat to the party.
27. **PP:** The man poked the monkey with a banana.
28. F: The cakes were chocolate.

Story 5: Beach

29. F: All these toys went to the beach.
30. **PP:** The boy saw the girl with binoculars.
31. F: There were 8 seashells.
32. F: This boy brought binoculars to the beach.
33. **RC:** The boy poked the girl using a pencil.
34. F: The fish were green.

Story 6: Zoo

35. F: All these toys went to the zoo.
36. **PP:** The man pointed to the girl with an umbrella.
37. F: This girl brought an umbrella to the zoo.
38. **RC:** The boy saw the girl using a magnifying glass.
39. F: There were 6 snakes.
40. F: The boy brought a magnifying glass to the zoo.

Story 7: Park

41. F: All these toys were at the park.
42. **RC:** The lady pointed to the man using a baseball bat.
43. F: This man brought a banana to the park.
44. F: There were 7 pieces of lime-flavored cake.
45. **PP:** The man poked the monkey with a banana.
46. F: The cakes were chocolate.

Story 8: Museum

47. F: All these toys went to the museum.
48. **PP:** The boy saw the girl with binoculars.
49. F: This boy brought a pencil to the museum.
50. F: This girl brought some binoculars to the museum.
51. **RC:** The boy poked the girl using a pencil.
52. F: There were 5 fish.

Story 9: Camping

53. F: All these toys went camping.
54. F: This girl brought the magnifying glass to the woods.
55. **RC:** The boy saw the girl using a magnifying glass.
56. F: There were 6 bugs.
57. **PP:** The man pointed to the girl with an umbrella.
58. F: The pigs were clean.

Appendix A2: Examples of story scripts

1. Teaching Story

This story is a circus story. In this story, there is this lady and this girl. And look, at the circus there is an elephant. So, the elephant is eating a strawberry and he says “Nomnomnom, strawberries are yummy!” Then the lady sees the elephant and she says “That elephant sure is cute.” But, the girl doesn’t notice, because she is going for a ride on this wagon, so she says, “Wee! I love riding in this wagon!”

Item	Target Response
The lady ate a strawberry.	False
The girl was riding in a wagon.	True
The girl watched the elephant eat a strawberry.	False
The lady watched the elephant eat a strawberry.	True

2. Story 4: Birthday Party Story

This is a birthday party story. And in this story, there is a doctor and it’s his birthday. There are also two monkeys, this man, and this lady. Everybody at the party brought a present for the doctor. The lady’s present was a baseball bat. The man and this monkey [monkey #1] brought a soccer ball. And, this monkey brought a banana for the doctor [monkey #2]. And look, because it’s his birthday, the doctor gets two chocolate cookies and two chocolate cakes.

So, the man with the soccer ball says “I love monkeys!” And he gives the monkey with the banana a big hug [monkey #2]. The lady with the baseball bat thinks this is funny, so she points her baseball bat at the man and goes “Heehee. It’s silly to hug a monkey.” Then, the doctor says “Yummmy! I love eating chocolate cakes and chocolate cookies on my birthday!”

Item	Target Response
F: All these toys were at a birthday party.	True
RC: The lady pointed to the man using a baseball bat.	True
F: There were 5 monkeys at the party.	False
F: This man brought a baseball bat to the party.	False
PP: The man poked the monkey with a banana.	False
F: The cakes were chocolate.	True

References

Abrams, K. H., Chiarello, C., Cress, K., Green, S., & Ellelt, N. (1978). The relation between mother-to-child speech and word order comprehension strategies in children. In R.N. Campbell & P.T. Smith (Eds.), *Recent advances in the psychology of language*. New York, NY: Plenum Press.

Altmann, G. (1986). *Reference and the resolution of local syntactic ambiguity: the effect of context during human sentence processing*. (Unpublished Doctoral Dissertation). University of Edinburgh, Edinburgh, U.K.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–238.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7 [Computer Software]. Retrieved from http://CRAN.R-project.org/package=lme4

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, *112*(4), 417-436.

Cairns, H. S., Waltzman, D., & Schlisselberg, G. (2004). Detecting the ambiguity of sentences: Relationship to early reading skill. *Communication Disorders Quarterly*, *25*(2), 68-78.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*(3), 687.

Crain, S. & Fodor, J. D. (1989). Competence and performance in child language. *Haskins Laboratories Status Report on Speech Research*, *SR-99/100*:118–36.

Crain, S., & McKee, C. (1985). The acquisition of structural restrictions on anaphora. In S. Berman, J.W. Choe, & J. McDonough (Eds.), *Proceedings of NELS 16*. Amherst, MA: University of Massachusetts.

Crain, S., Thornton, R., Boster, C., Conway, L., Lillo-Martin, D., & Woodams, E. (1996). Quantification without qualification. *Language Acquisition, 5*(2), 83-153.

Felser, C., Marinis, T., & Clahsen, H. (2003). Children's processing of ambiguous sentences: A study of relative clause attachment. *Language Acquisition*, *11*(3), 127-163.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies.* (Unpublished Doctoral Dissertation). University of Connecticut, Storrs, CT.

Gillis, R., & Nilsen, E. S. (2014). Cognitive flexibility supports preschoolers' detection of communicative ambiguity. *First Language*, *34*(1), 58-71.

Gordon, P. (1996). The truth-value judgment task. In D. McDaniel, C. McKee, & H. Cairns (Eds.), *Methods for assessing children's syntax*. Cambridge, MA: MIT Press.

Gordon, P., & Chafetz, J. (1986). *Lexical learning and generalization in the passive acquisition*. Paper presented at the 11th Annual Boston University Conference on Language Development, Boston, MA.

Hamburger, H. (1980). A deletion ahead of its time. *Cognition, 8*(4), 389-416.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L., & Trueswell, J. (2001). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research, 29*, 597-626.

Johnston, J. R., & Slobin, D. I. (1979). The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language*, *6*(3), 529-545.

Kidd, E., & Bavin, E. (2005). Lexical and referential cues to sentence interpretation: An investigation of children's interpretations of ambiguous sentences. *Journal of Child Language*, *32*(4), 855-876.

Kidd, E., & Bavin, E. (2007). Lexical and referential influences on online spoken language comprehension: A comparison of adults and primary-school-age children. *First Language*, *27*(1), 29-52.

Kidd, E., Stewart, A., & Serratrice, L. (2011). Children do not overcome lexical biases where adults do: The role of the referential scene in garden-path recovery. *Journal of Child Language*, *38*(1), 222-234.

Limber, J. (1973). The genesis of complex sentences. In T. Moore (Ed.), *Cognitive Development and the Acquisition of Language*. New York, NY: Academic Press.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676-703.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake--but only for skilled producers. *Journal of Experimental Psychology*, *38*(4), 843-847.

McKee, C., McDaniel, D., & Snedeker, J. (1998). Relatives children say. *Journal of Psycholinguistic Research, 5*, 573–595.

Omaki, A., & Lidz, J. (2015). Linking parser development to acquisition of syntactic knowledge. *Language Acquisition*, *22*(2), 158-192.

Peterson, C., Dowden, C., & Tobin, J. (1999). Interviewing preschoolers: Comparisons of yes/no and wh-questions. *Law and Human Behavior*, *23,* 539–555.

R Core Team. (2014). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, *49*, 238-299.

Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language*, *58*(2), 574-608.

Spivey-Knowlton, M.J., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, *55*, 227-267.

Steffensen, M.S. (1978). Satisfying inquisitive adults: Some simple methods of answering yes-no questions. *Journal of Child Language, 5*, 221–236.

Tomasello, M. (1987). Learning to use prepositions: A case study. *Journal of Child Language*, *14*(1), 79-98.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying online sentence processing in young children. *Cognition*, *73*(2), 89-134.

Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In C. Clifton Jr., L. Frazier & K. Rayner (Eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Erlbaum.

Trueswell, J. C., & Tanenhaus, M. K. (2005). *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*. Cambridge, MA: MIT Press.

Wankoff, L. S. (1983). Selected metalinguistic variables as they relate to conservation and reading achievement in normal achieving and some learning-disabled children. (Unpublished Doctoral Dissertation). City University of New York, New York, NY.

Valian, V. (1986). Syntactic categories in the speech of young children. *Developmental Psychology*, *22*(4), 562-579.

Yuill, N. (2009). The relation between ambiguity understanding and metalinguistic discussion of joking riddles in good and poor comprehenders: Potential for intervention and possible processes of change. *First Language*, *29*(1), 65-79.

Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika*,*70*(1), 203-212

Appendix B. Children's Awareness of Syntactic Ambiguity (Article 2)

Abstract

This study probes children's ability to consciously notice when a sentence has multiple meanings. Previous research suggests that children do not detect syntactic ambiguity until second grade, but once this skill develops, it aids in reading comprehension. The current study asks whether younger children can detect syntactic ambiguity using a picture selection and explanation task. Participants included 16 adults and 55 children (aged 4;10 to 7;0). To test for conscious awareness, children were encouraged to teach a confused puppet why multiple pictures could match one sentence. About 16% of the children showed a strong awareness of syntactic ambiguity by giving adult-like explanations on at least half of the experimental items. Furthermore, there was a trend for children's explanations of the ambiguity to improve with age. The variation in children's performances both on picture selections and explanations of the ambiguity suggests that the ability to notice syntactic ambiguity is a complex skill that develops over time.

Keywords: syntactic ambiguity detection, metalinguistic awareness, children, language development

1.0 Introduction

Metalinguistic awareness is the ability to consciously think about language as an object (Cazden, 1974; Demont & Gombert, 1996). It begins in pre-school, with certain metalinguistic skills developing earlier than others (Bialystok, 2001; Ehri, Nunes, Stahl, & Willows, 2001). Some metalinguistic skills are precursors to literacy development. For example, phonological awareness supports children's ability to decode letters in the orthography to sounds in the language (Ehri et al., 2001). As decoding is a critical part of early reading development, it is no surprise that early phonological awareness correlates with future reading ability (Goswami & Bryant, 1990). While there has been considerable research on the connection between phonological awareness and reading (see Goswami & Bryant, 1990; Castles & Coltheart, 2004 for review papers), somewhat less is known about how syntactic awareness connects to reading development (Tong, Deacon, & Cain, 2014).

There are both empirical and theoretical reasons to think that syntactic awareness contributes to reading. The well-known model of reading, the Simple View of Reading (SVR), suggests that two components determine reading comprehension: decoding and linguistic comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990; Joshi & Aaron, 2000). Linguistic comprehension is the ability to derive an interpretation from a sentence and integrate it in the discourse. Helen Smith Cairns and her colleagues suggest that syntactic awareness contributes to linguistic comprehension, and thus, should contribute to reading development (Cairns, Waltzman, & Schlisselberg, 2004; Cairns, 2013). Empirical research supports this. For example, several studies have shown that grammatical awareness, which is the ability to detect and repair ungrammatical utterances, correlates with successful reading (Bentin, Deutsch, & Liberman, 1990; Demont & Gombert, 1996; Nation & Snowling, 2000; Tong et al., 2014). Another component of syntactic awareness is syntactic ambiguity detection, or the ability to consciously recognize when a structurally ambiguous sentence has multiple meanings. Some research suggests that this ability also contributes to reading development. Cairns et al. (2004) found that syntactic ambiguity detection in second graders correlated with their reading scores in the third grade. Yuill (2009) and Zipke, Ehri, & Cairns (2009) found that 7- to 9-year-olds' improvement at syntactic ambiguity awareness correlated with improved reading comprehension. This connection to reading development makes it important that we understand the development of syntactic ambiguity awareness.

Research so far suggests that children do not spontaneously detect syntactic ambiguity until the second grade, which is around 8-years-old (Wankoff, 1983; Cairns et al., 2004). Earlier research had claimed that the skill emerges as late as 12-years-old (Shultz & Pilon, 1973; Keil, 1980). The current study probes 4- to 7-year-olds' syntactic ambiguity detection, asking whether some children begin to develop the ability well before second grade. If some younger children can detect syntactic ambiguity, its contribution to reading development may begin earlier than 8-years-old. It would be important to discover this as ambiguity detection may be a skill that should be incorporated into early literacy curriculum.

I hypothesize that some children under 8-year-olds will show awareness of syntactic ambiguity. One reason to expect that comes from work on other types of syntactic awareness. Like syntactic ambiguity detection, the ability to judge well-formedness of sentences is a type of metalinguistic awareness related to syntax. It was thought for many years that this ability did not develop until first or second grade (Scholl & Ryan, 1980; Tunmer, Pratt, & Herriman, 1984). However, Dana McDaniel and colleagues have shown that children as young as 3-years-old can judge a variety of sentence types if it is clear to them that they are discussing language as an object (McDaniel & Cairns, 1990; 1996; Cairns, Schlisselberg, Waltzman, & McDaniel, 2006). Well-formedness judgments may differ from syntactic ambiguity detection, as the former tends to require binary judgments while the latter generally requires a child to discuss language. Nevertheless, this work establishes that young children do have metalinguistic awareness of their syntactic system, suggesting that some may be able to detect syntactic ambiguity.

I also predict that there will be a great deal of variation in children's detection of syntactic ambiguity. Even adults are quite varied on metalinguistic abilities (Gleitman & Gleitman, 1971; Birdsong, 1989). Many factors can influence a person's metalinguistic skills and contribute to this variation. For example, multilinguals outperform monolinguals on many metalinguistic skills (Aronsson, 1981; Bialystok, 1986; Hermanto, Moreno, & Bialystok, 2012). There are also correlations between children's metalinguistic awareness and their proficiency at conservation, which is understanding that an amount stays the same when it changes container or shape (Hakes, 1980; Schlisselberg, 1988; van Kleeck, 1994). Furthermore, it is unlikely that children develop syntactic ambiguity awareness all at once. Metalinguistic skills likely involve complex relations between the linguistic system and metacognition (Hakes, 1980; Bialystok & Ryan, 1985). Syntactic ambiguity detection has at least two components. First, a person must access both interpretations of an ambiguous sentence. Then, a person must become consciously aware of the ambiguity. Beyond this, it also requires a certain level of verbal skills to articulate why a sentence is ambiguous. It is not necessarily the case that a person must be able to explain an ambiguity in order to have awareness of it. However, methodologically, our most direct evidence of a person's conscious awareness of ambiguity comes from their explanations of it (Cairns et al., 2004; Wankoff & Cairns, 2009). Because syntactic ambiguity detection requires several components, it is likely that this ability develops over time. Because of the considerable variation in children's metalinguistic skills, it is also likely that syntactic ambiguity awareness develops in children at different times and rates. Thus, in the current study, I predict that some children will already have awareness of syntactic ambiguity, while others will only be beginning to develop the ability. Some children may not show any indications of syntactic ambiguity awareness yet.

It is important, though, to address why children under 8-years-old may have failed at syntactic ambiguity detection tasks in previous research. I see three possible explanations. The first is that the type of metalinguistic awareness needed to detect syntactic ambiguity does not develop until 8-years-old. This is what most research has concluded (Shultz & Pilon, 1973; Hirsh-Pasek, Gleitman, & Gleitman, 1978; Keil, 1980; Wankoff, 1983). This explanation attributes failure at such detection to differences in children's and adults' linguistic metacognition, not differences in their linguistic system per se. That is, both children and adults can access both interpretations of a syntactic ambiguity, but only adults can consciously consider both meanings. Another possible explanation relates to the developing parser. One way in which children's parsing differs from that of adults is that children are less likely to revise an initial parse than adults are (Trueswell, Sekerina, Hill, & Logrip, 1999; Snedeker & Trueswell, 2004; Kidd, Stewart, & Serratrice, 2011; Omaki & Lidz, 2015). Appealing to this, Cairns (2013) argued that children may not detect syntactic ambiguity because they do not parse both interpretations in the first place. This explanation relies on differences in the child and adult's parsers and not on their metacognition. However, in Zimmer (2016), I used a truth value judgement task to show that 3- to 5-year-olds usually accept an ambiguous utterance as true in different discourse contexts that support the different interpretations of the utterance. Crucially, the study included contexts where the same utterances were untrue. Children generally accepted an utterance as true for both contexts supporting its two interpretations but rejected it in a context where it was false. This hints that children can and do access both interpretations of a syntactic ambiguity. Questions regarding the timeline of such access of course remain.

The third possible explanation for children's failure to detect syntactic ambiguity relates to the methods used in most previous research on this skill (Shultz & Pilon, 1973; Keil, 1980; Wankoff, 1983; Cairns et al., 2004). These studies' focus was spontaneous ambiguity detection. Thus, they all used interview tasks in which children were asked explicitly if a sentence can mean more than one thing (Shultz & Pilon, 1973; Keil, 1980; Wankoff, 1983; Cairns et al., 2004). For example, in Cairns et al. (2004), children first heard a sentence and were asked if it could mean two things. If this question failed to elicit an interpretable description, children were prompted verbally and then with pictures representing both interpretations. Children's responses were scored on a weighted system based on whether they noticed the ambiguity spontaneously, after the verbal prompt, or after the picture prompt. Scores for 4- to 7-years-olds were very low. It was not until second grade that some children's responses showed spontaneous detection of syntactic ambiguity. However, it may be that this interview task's emphasis on spontaneous detection caused it to miss early indications of syntactic ambiguity awareness. As previously mentioned, ambiguity detection is likely not a monolithic skill that develops all at once. It is more likely that aspects of it develop before a child can spontaneously detect and describe a syntactic ambiguity simply upon hearing it.

Thus, the current study used a picture selection task that I designed to test for conscious awareness of syntactic ambiguity. Participants first chose from a set of four pictures all the ones that they thought matched an utterance. For ambiguous items, two pictures could match. For unambiguous items, only one picture could match. Then, participants explained to a confused puppet why they chose those pictures. This task had several strengths. The picture selection portion provided visual support for the ambiguity and also was engaging for children (Gerken & Shady, 1996), while the puppet made children more comfortable teaching (Thornton, 1996). Furthermore, having both picture selections and explanations provided a more nuanced measure for detecting early awareness of syntactic ambiguity. Children's picture selections alone provided weak evidence for their ambiguity detection. For children without the verbal skills to explain why a sentence is ambiguous, their picture choices may reveal early signs of ambiguity awareness. However, our interpretation of the picture selections should be limited because there are other reasons that a child could choose both target pictures than having detected the ambiguity. For example, the two target pictures looked more similar to each other than the distractor pictures. Furthermore, on the adjective ambiguity items, because the first noun phrase is modified in both interpretations, a participant could theoretically choose both target pictures having only computed the interpretation where the first noun phrase is modified. Therefore, the stronger and more reliable evidence for ambiguity awareness comes from participants' explanations of their choices. The scoring system developed for the study allows for comparison of children's explanations to adult explanations. Thus, it is able to measure gradient differences in children's ability to explain syntactic ambiguity and thus indicate a conscious awareness of the ambiguity.

1.1 Ambiguity types

This study includes two types of syntactic ambiguity. The first occurs when a conjoined noun phrase (NP) is modified by a prenominal adjective, as shown in (1) and (2). I refer to this as an adjective ambiguity, shortened to Adj in some tables.

(1) The girl is noticing [dirty [pigs and ducks]].[1]
(2) The girl is noticing [[dirty pigs] and ducks].

In (1), the prenominal adjective *dirty* modifies the entire conjoined NP so the interpretation is that both the pigs and ducks are dirty. In (2), *dirty* only modifies the first noun, so the interpretation is that only the pigs are dirty and the ducks could be any state.

The second ambiguity type occurs when a prepositional phrase (PP) at the end of a transitive sentence can attach either to the local NP or to the verb phrase (VP), as shown in (3) and (4). I refer to this as a PP ambiguity.

(3) The man [is pointing [to the girl with the umbrella]].
(4) The man [is pointing [to the girl] with the umbrella].

When the PP attaches to the local NP as in (3), the interpretation is that the man is pointing to the girl who has the umbrella. I refer to this as the NP-attachment interpretation. When the PP

---

[1] The perceptual verb *notice* is not felicitous for some people in the progressive tense. Several young adults (ages 25 and younger) from the location of the study reported that it was acceptable to them. Pilot work with children of the same age and younger than those in this study accepted my use of this verb to match the relevant pictures.

attaches to the VP as in (4), the interpretation is that the man is using the umbrella to point to the girl. I refer to this as the VP-attachment interpretation.

I included PP ambiguity because it is the subject of most research on children's comprehension and processing of syntactic ambiguity to date. Thus, we know more about how children handle this type of ambiguity than any other. I included adjective ambiguity because, in my own pilot work, I found that children were more likely to accept both meanings of this type of ambiguity than PP ambiguities in a truth value judgement task. Thus, this type of ambiguity may be more salient to children than PP ambiguities.

2.0 Methods

For every experimental sentence, participants were shown four pictures on a game board. An experimenter instructed participants to choose all of the pictures on the board that matched that sentence. Again, when an item was ambiguous, two pictures corresponded to the two meanings of the sentence. When an item was unambiguous, only one picture on the board corresponded to it. So the target choices were always one or two pictures from the board. After selecting pictures, participants explained their choices to a puppet that was controlled by a second experimenter. We recorded and later transcribed and coded these explanations. The task was introduced with a teaching dialogue that explicitly informed participants that language can be ambiguous with lexical and sentential examples. Children were taught the task with three types of teaching items: a lexical ambiguity, a syntactically ambiguous sentence, and an unambiguous sentence.

2.1 Participants

Participants included 55 children ranging in age from 4;10 to 7;0 and 16 adults. All participants were native English speakers. All adult participants were monolingual; 15 child participants were bilingual.[2] Based on parent report, no participants had any language-related disorders. The children's age range matters because previous research (e.g., Cairns et al., 2004, pp. 73) specifically claims that children this young cannot detect syntactic ambiguity. A total of 62 children were tested. Three were excluded from analyses because a recorder malfunction caused part of their data to be deleted. Four children were excluded because they failed the inclusion criteria. Specifically, they chose pictures that were unrelated to the utterance on three or more of eight items. I took this pattern to suggest that those children had not learned the task. Child participants were tested at private and charter preschools and elementary schools in

[2] Nine of the bilingual children were speakers of Spanish. The other six bilingual children spoke French, German, Hebrew, Tamil, Mandarin, and Vietnamese. No information was collected about their degree of bilingualism. Ten of the bilingual participants are enrolled in a school with a focus on bilingual immersion.

Tucson, Arizona.[3] Adult participants were undergraduate students at the University of Arizona.[4] No adults failed the inclusion criteria.

2.2 Teaching dialogue

A key element of the experimental design was the teaching dialogue. This dialogue taught the task, but it was not intended to provide training on ambiguity detection per se. Before describing the task, we established language as the topic of conversation by introducing ourselves as scientists who study language and then demonstrating a lexical ambiguity. The child was asked to label two pictures representing the two different meanings of the word *bat* (i.e., a baseball bat and a flying mammal). The child's attention was drawn to the fact that he or she used the same word for two different pictures, and that meant that the word *bat* means two different things. The first teaching item used lexical ambiguity because previous research has shown that children are more aware of lexical ambiguity than syntactic ambiguity (Cairns et al., 2004; Wankoff & Cairns, 2009). Thus, the task itself was taught with an ambiguity that the child was likely to notice. After establishing language as the topic, one experimenter introduced the child and the puppet to the picture game, explaining that words and sentences with two meanings match two pictures and sentences with one meaning match only one picture.

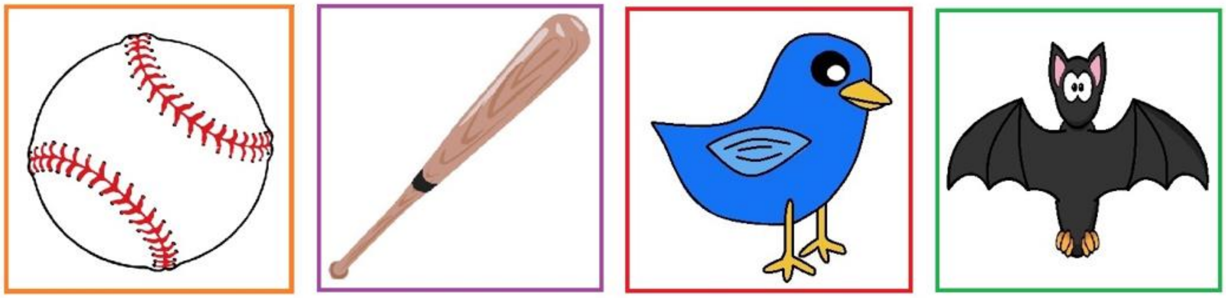
2.3 Teaching materials

Three types of teaching items made up the teaching sequence: an ambiguous word, an ambiguous sentence, and an unambiguous sentence. The first such item used the word *bat*, which the child had just discussed. Children were shown two pictures representing the two meanings of *bat*, along with two distractor pictures, as shown in Figure 1. Each distractor image was semantically related to one of the target pictures (i.e., *bird* to the animal bat, *ball* to the baseball bat). If the child did not select both pictures representing the two meanings of *bat*, the child was given feedback by the experimenters. Then, the child did a back-up lexical ambiguity item, *nail*, and the sequence was repeated. A third lexical ambiguity item, *star*, was prepared, but no participant failed to choose both pictures for both *bat* and *nail*.

---

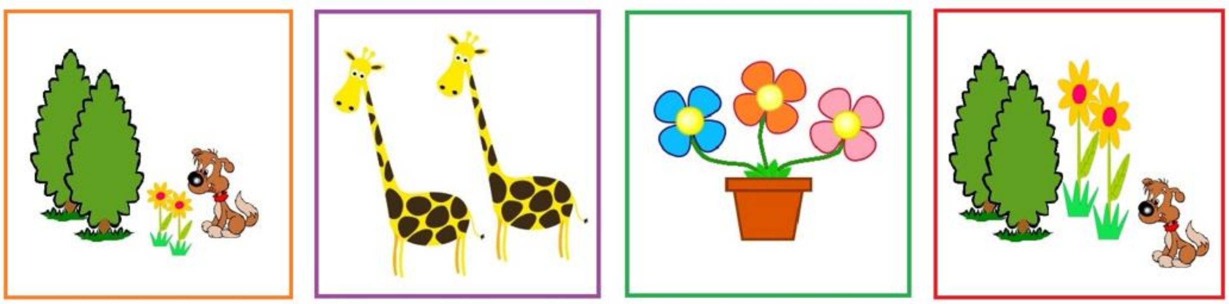
[3] Private and charter schools were chosen because of time constraints for the study. It can take several months to gain permission to conduct research in Tucson-area school districts. The schools that participated in this research had student populations from relatively high income families as compared to the rest of Tucson.
[4] Adult participants were not systematically asked if they had knowledge of linguistics. However, none reported having studied linguistics during an informal debriefing session.

Figure 1: *bat*



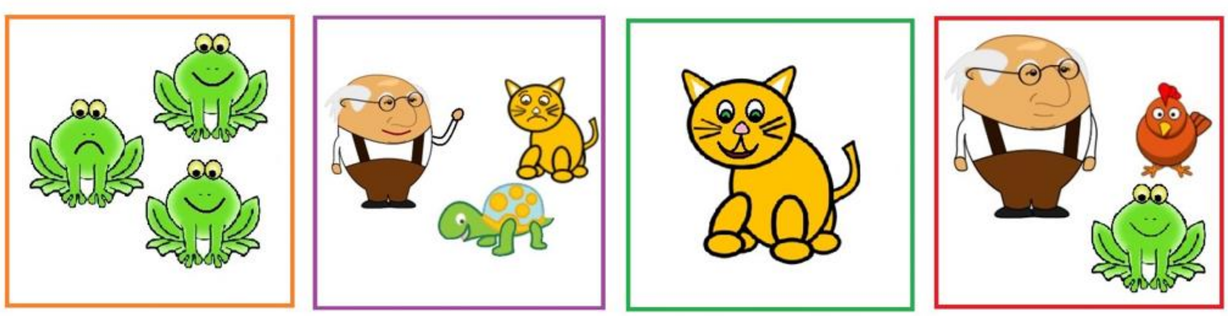
The next teaching item was an ambiguous sentence. Children saw two pictures representing the two interpretations of the sentence along with two distractor pictures, as shown in Figure 2. One back-up ambiguous sentence teaching item was prepared (*The man is poking the monkey with banana.)* If a participant failed to choose both pictures, experimenters gave feedback and then did the back-up item.

Figure 2: *The puppy is sniffing tall bushes and flowers.*



Children were then presented with an unambiguous sentence. They saw one picture that matched the sentence and three distractor pictures, as shown in Figure 3. One back-up unambiguous sentence teaching item was prepared (*The boy is watching the girl, and she has a magnifying glass).* If a participant chose more than one picture, experimenters did the back-up item. Once the teaching items were complete, participants continued on to experimental items.

Figure 3: *The grandpa is waving at the happy turtle and the sad cat.*



2.4 Experimental materials

The four ambiguous experimental sentences included two PP ambiguities and two adjective ambiguities. There were also four unambiguous experimental sentences. As Table 1 shows, unambiguous sentences mirrored the structure of ambiguous sentences. Two unambiguous sentences mirrored the PP ambiguities, and two mirrored the adjective ambiguities.

Table 1: Experimental items

Ambiguous Sentences		**Unambiguous Sentences**
PP	1. The man is pointing to the kid with an umbrella. 2. The boy is staring at the girl with binoculars.	1. The grandma is looking at the bear, and it has a telescope. 2. The woman is using a feather to tickle the baby.
Adj	1. The teacher is studying green frogs and fish. 2. The girl is noticing dirty pigs and ducks.	1. The doctor is wanting vanilla cakes and chocolate cookies. 2. The chef is making strawberry cupcakes and strawberry ice cream.

For each ambiguous experimental sentence, two pictures represented the two interpretations of the sentence. These were shown with two distractor images. For each unambiguous experimental sentence, one picture represented its meaning. It was shown with three distractor images. For more regarding the distractor images, see Appendix A.

2.5 Data treatment

Participants' picture selections were written down by both experimenters and later coded as expected or not expected responses. The expected response for unambiguous items was choosing only the picture representing the one interpretation of the sentence. The expected response for ambiguous items was choosing all and only the two pictures representing the sentence's two interpretations.

All experimental sessions were audio recorded. Using the recordings, participants' explanations of their picture choices were transcribed. For every ambiguous item, responses were scored on a scale from 0-3. This scale was developed based on earlier testing of adults. It reflects regular properties of their explanations of ambiguities. Previous studies have not included scoring systems based on specific properties of adult explanations. This is important though if we see adult metalinguistic skills as the child's target. Specifically, explanations from adults typically described an explicit difference and similarity between the two interpretations. By explicit, I mean that adults mentioned specific elements of the images that corresponded to the relevant concepts in the ambiguous sentence. For example, rather than saying "the pictures look similar," adults would say "the pigs are dirty in both pictures." Examples of adult explanations that include an explicit mention of a similarity and a difference are in (5) and (6); these both exemplify what was coded as a 3 response in the children's explanations.

(5) "First here's the boy and there's the girl and he's staring at the girl in both and then 'with binoculars' could either mean he's staring at the girl with them or she has the binoculars."
(6) "So the sentence could mean the girl is noticing dirty pigs and dirty ducks or it could mean the girl is noticing dirty pigs and just ducks."

Thus, the high end of the scoring scale, shown below in Table 2, advantaged explanations that referenced both an explicit similarity and an explicit difference. The table also shows examples of child explanations at each level of the scale (ages of speakers in parentheses). Note that, because of the explicitness requirement, explanations that only mentioned that two pictures were similar or different without mentioning specific elements of the images received a score of 1. Because there were four experimental utterances, an individual's maximum score was 12.

Table 2: Explanation scoring system.

Score	**Picture Choice**	**Participant Explanations**	**Examples of Child Explanations**
0	only one target or distractors	N/A	N/A
1	both targets	discussed choices without explicit similarity or difference	"Because I picked two." (5;6)
			"Because they match." (6;2)
2	both targets	discussed either explicit difference or similarity	"They're the only ones with binoculars and the boy and the girl." (6;11)
			"Both of them have fish and frogs." (6;7)
3	both targets	discussed both explicit difference and similarity	"Because look it can be like he's pointing with an umbrella to the kid or he's pointing to the kid with an umbrella. Even though the pictures are different there's still an umbrella, a man, and a kid." (6;4)
			"Because this is the one where he points it with his finger and she has the umbrella and this one is where he's pointing the umbrella to her." (6;7)

To validate this scoring system, a group of 16 adult control participants did the study. Explanations from the adult control group scored as follows: 45/64 of the total ambiguous items scored a 3; 8/64 scored a 2, and none scored a 1. The other 11 items were cases of the participant not choosing both pictures in the first place because he or she failed to notice the ambiguity, scoring a 0. Thus, when adults noticed ambiguity, their explanations mentioned an explicit similarity and difference between the interpretations about 85% of the time.

As the examples in Table 2 show, children's explanations range from very simple to quite sophisticated. The 3-point explanations very much resemble the adult explanations in (5) and (6). Thus, this scoring system can measure for gradient improvements in children's explanations of syntactic ambiguity. However, this scoring system has certain weaknesses as well. It may underrepresent the number of children who noticed the ambiguity. For instance, some children may notice the ambiguity but not have the verbal skills to articulate it. The picture selection portion was included to allay this issue. Children who did not articulate the ambiguity well could still show a weak indication of ambiguity awareness by choosing both pictures. Additionally, because the scoring system only focuses on two properties of adult explanations, it may miss other properties that indicate ambiguity awareness.

For example, another interesting feature of adult explanations of adjective ambiguities was mention of the adjective's underspecificity. In other words, adults often pointed out that it was not specified whether the adjective applied to the second noun or not. Examples of adult explanations with this feature are in (7) and (8).

(7) "Well both of the frogs are green and there wasn't necessarily specified what color the fish were so that's why I chose both cause it could be orange or it could be green."
(8) "Because the color of the fish wasn't said so I put both of them because it could be either."

In the adult control group, eight out of 16 adults gave at least one explanation that exhibited this feature. As I will discuss in greater detail below, some children produced this type of explanation as well. Note that it requires metalinguistic sophistication. I did not include this property in the scoring system though because it can only apply to the adjective ambiguity type. There are likely to be other such properties specific to certain ambiguity types. In future research, a more nuanced coding system may find more indications of children's ambiguity awareness. The following section describes the results of my scoring system applied to all participant's explanations.

3.0 Results and Analyses

In this section, I will report on three types of data for adults, younger children (4-5-year-olds) and older children (6-7-year-olds). Later, for the purposes of statistical analyses, I will collapse the data across children. The first data type is participants' percentage of correct picture selections. For ambiguous items, this meant choosing both pictures representing the sentence's two interpretations and no others. For unambiguous items, this meant choosing only the one picture that represented the sentence's interpretation. Including both unambiguous and ambiguous sentences, the adult mean for correct picture selection was 91% (83% for ambiguous items and 98% for unambiguous items). The younger children's mean for correct picture selection was 56% (30% for ambiguous items and 81% for unambiguous items). The older children's mean for correct picture selection was 60% (24% for ambiguous items and 96% for unambiguous items).

The second data type is participants' total scores, which is the sum of a participant's scores on all ambiguous items and is on a scale from 0-12. The third type of data, which I discuss later, is individual item scores which are on a scale from 0-3. As expected, the data shows considerable variation in the child group and to a lesser extent in the adult group. Overall, adults scored much higher than the children.[5] Table 3 compares the score ranges, medians, and means. Standard deviations are not included because the distributions were skewed for all groups. As expected, adults were skewed towards ceiling. Both child groups were skewed towards floor. Even though the means for the children were low, some individual children in both age groups scored as high as 9 (which is like the adult mean).

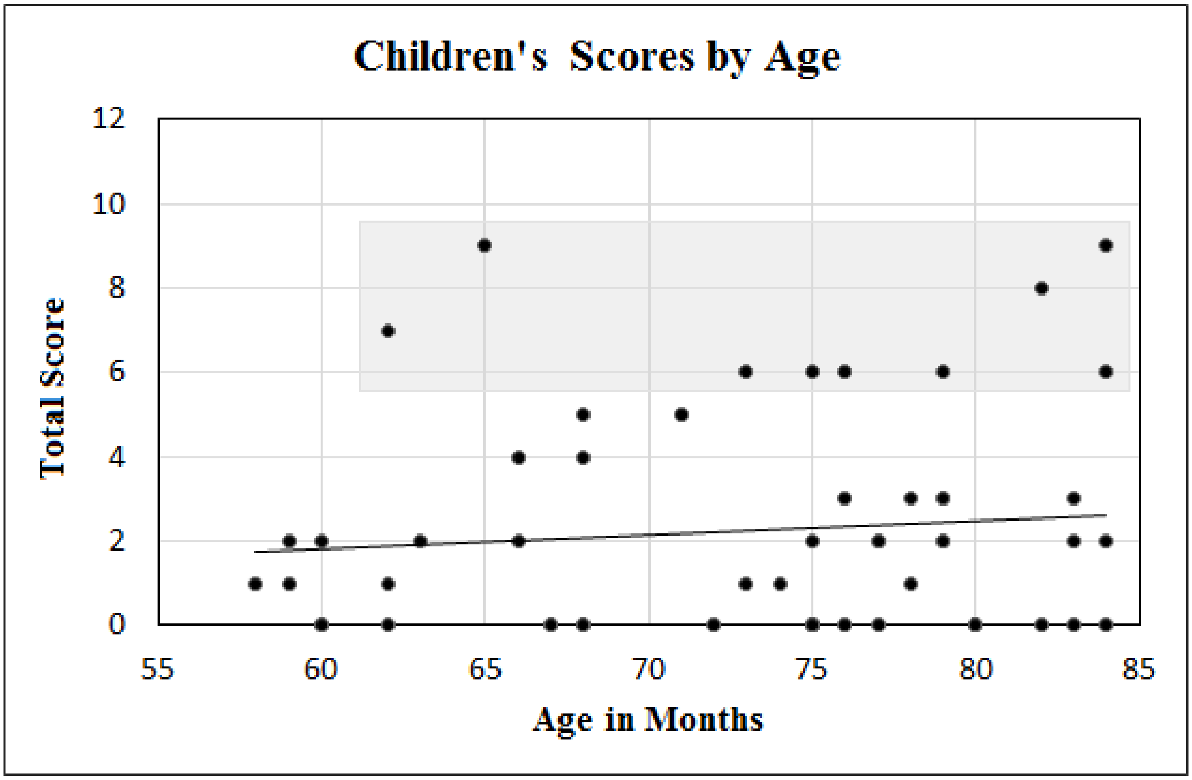
Table 3: Average total scores by group (scores out of 12).

Age Group	**# of Participants**	**Score Range**	**Mean**	**Median**
Adults	16	3-12	9.4	10.5
Older Children	35	0-9	2.3	2
Younger Children	20	0-9	2.3	1.5

Figure 4 plots children's total scores by their age in months. From this scatter plot, we can see that nine children scored 6 or higher (indicated by the gray box). In other words, these children produced adult-like responses on at least two of the four experimental items. Two of these were 5-year-olds. The rest were between 6;1 and 7;0.

[5] Because bilingualism is known to affect metalinguistic skills (e.g., Bialystok, 1986; 2001), I also compared the mean total scores for the 15 bilingual children and the 40 monolingual children separately. The bilingual group had a higher mean at 3.8 and the monolingual group had a lower mean at 1.9. While it seems that the bilinguals had an advantage in this experiment, there is a confound: Their mean age was six months older than that of the monolinguals. Future research should systematically compare matched monolinguals to bilinguals to probe whether bilingualism improves children's awareness of syntactic ambiguity.

Figure 4: Children's total scores plotted by age in months.



While more 6-year-olds than 5-year-olds were in this group of high performers, a Kendall's Tau Rank correlation test, which is non-parametric, revealed that children's total scores were not significantly correlated with their age in months (z=0.69, p>.05).

As this scatter plot shows, age did not predict total scores in the 55 children plotted here. However, the representation in Figure 4 is somewhat skewed by the 21 children who received scores of zero (ranging in age from 5;4 to 7;0). Each of these children chose only one picture for every experimental item. There are a few possible explanations for why an individual child would always select one picture. As Cairns (2013) suggests, it could be that such a child was unable to revise his or her initial parse of each sentence and thus did not notice the ambiguity. Another possibility is methodological. Some children may have adopted a response strategy for the game, always selecting one picture. Experimenter feedback during the teaching dialogue was meant to offset this possibility; however, it is still possible that some children developed such a strategy. Examples of these children's explanations suggest that there were both types of children in the study.

First, consider an explanation that suggests the child was not revising initial parses. In this example, the child was explaining why a picture did not match the sentence *The man is pointing to the girl with an umbrella.* She chose the image where the man was using the umbrella to point to the girl (VP-attachment). In the comment in (9), the child was explaining why the other target picture (i.e., a girl holding an umbrella) did not match the sentence.

(9) “So this one is pointing at a girl with an umbrella so that doesn't match.” (5;8)

Interestingly, this child repeated a part of the experimental sentence almost verbatim when describing the picture that she claimed did not match the sentence. This suggests she was so strongly biased towards the VP-attachment interpretation of the sentence that she was unable to accept the NP-attachment interpretation even as she produced almost the exact same sentence when describing the NP-attachment image. Two other children gave similar explanations throughout the study. This suggests that some children only chose one picture because they only accessed one interpretation of the sentence. This fits with an explanation where the components necessary for syntactic ambiguity detection develop at different times in different children. While some children in this study are already showing a strong awareness of ambiguity, others might not have even accessed both interpretations at the same time. Between these ends of the spectrum, there are children who often chose both pictures but did not give adult-like explanations. This suggests that syntactic ambiguity detection is a complex skill. It requires the ability to access both interpretations, which is revealed by the picture selection task, and the ability to notice and describe the ambiguity, which is revealed by the explanation task.

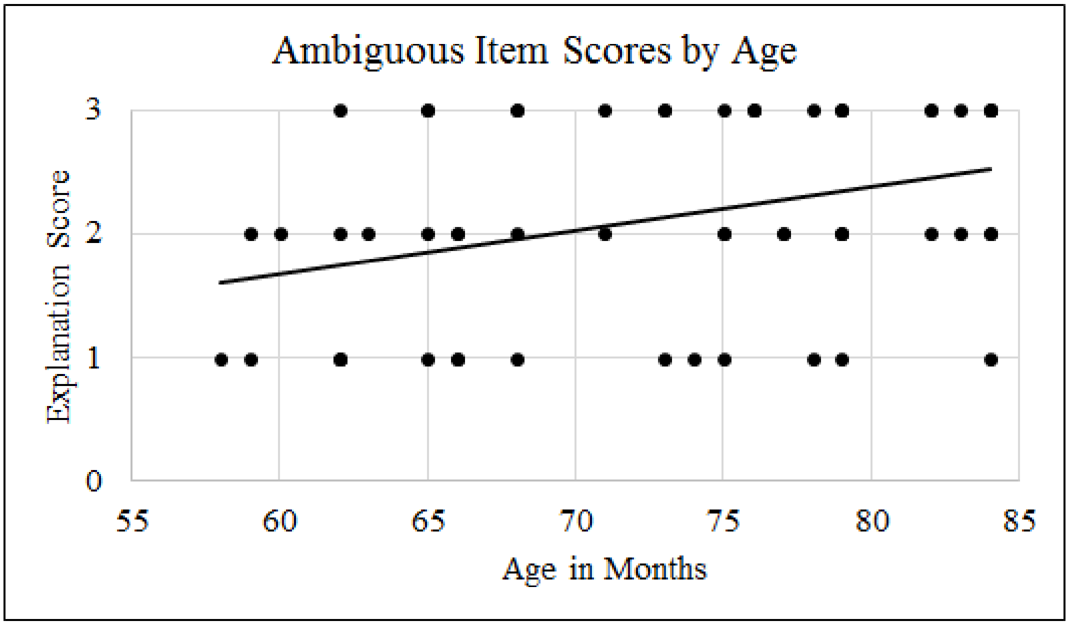
On the other hand, some children who always chose one picture responded in ways that suggest they had adopted a response strategy in the experiment. For example, five children gave roughly the same explanation for three or four of the four ambiguous items. In each case, these explanations were undetailed, as exemplified in (10) and (11).

(10) “because it matches the sentences” (6;8)
(11) “because there's one of them” (4;10)

The fact that these children gave the same undetailed explanation for their choices for most items suggests that they may have adopted a strategy for the game. Therefore, it seems as though the children who always chose only one picture were varied amongst themselves. Some gave clear explanations suggesting that they truly failed to access the other interpretation of the sentence. Others seemed to have adopted a strategy to only choose one picture for each sentence.

Children who scored 0 may have skewed the data, washing out any developmental trend in Figure 4 above. However, there is another possible relationship between age and children’s performance. It is possible that age may not affect picture choices but it might affect the quality of children’s explanations when they do choose both pictures. To examine this relationship, Figure 5 shows the scores for explanations of individual items, only including those where the child chose both pictures.

Figure 5: Ambiguous item explanation scores plotted by age in months.



When only items such items were included, a Kendall's Tau Rank test revealed that children's individual explanation scores were positively correlated with their age in months (z=2.59, p<.01).

Figures 4 and 5 show that older children were not more likely than the younger children to choose both pictures on ambiguous items, but they were more likely to give adult-like explanations when they did select both pictures. Thus, there is a relationship between age and quality of explanation but not picture choices.

Finally, four children (ages 6;6-7;1) gave explanations that exhibited the adult strategy of mentioning the underspecification of the adjective in their explanations of adjective ambiguities. Examples of such explanations are shown in (12) and (13).

(12) "Because it doesn't matter which color cause these are fishes and this is a frog and these are fishes and these are frogs." (6;11)

(13) "This has two fish and this one has green fish. […] Even though the fish are not the same color, they're still fish." (6;4)

This type of explanation of ambiguity has not been discussed in previous research on children's ambiguity awareness. However, such explanations suggest some metalinguistic sophistication as they involve an explicit acknowledgement that the adjective only specified the property of one of the nouns. In the scoring system, this was treated as a mention of a difference because it is acknowledging that the two nouns have different properties. Interestingly, three of the four children gave this type of explanation for the only ambiguous item for which they selected both pictures. Thus, these four children had total scores of 2, 2, 3, and 6.

3.1 Group comparisons

To compare adults to children, I used generalized linear mixed effect logistic models (GLMEs) in the R lme4 package (R Core Team, 2014; Bates, Maechler, Bolker, & Walker, 2014). Random intercepts included item and subject in every model. Item was included as a random intercept because the items were presented in the same order for all participants. Thus, the models can account for ordering effects. The models had no random slopes. For the first set of analyses, the dependent measure was expected/non-expected response. I chose GLMEs for this data because they can specify a binomial distribution for a binary dependent measure (Baayen, 2008; Winter, 2013). The first model compared children's and adults' numbers of expected responses on only the ambiguous items. Fixed effects included group (adults, children), ambiguity type (PP, Adj), and the interaction of the two. The interaction was not significant (Est.=-1.07, SE= 0.86, z=-1.25, p>.05). Because of this, two models were fitted with only group as the fixed effect and only ambiguity type as the fixed effect to calculate the main effects. The main effect of group was significant with adults giving the expected response more than children (Est.=-3.83, SE=0.72, z=-5.29, p<.001). The main effect of ambiguity type was not significant (Est.=-1.20, SE=0.64, z=--1.87, p>.05).

GLMEs were also used for analyses of participants' coded explanations, scored from 1-3. These only included items where participants selected both pictures. The dependent measure was the scores assigned to participants' explanations. GLMEs were chosen here because they can specify a Poisson distribution for categorical data (Baayen, 2008; Winter, 2013). Again, subject and item were included as random intercepts in all models, and there were no random slopes. The first model compared children's and adults' explanation scores by ambiguity type. Fixed effects included group (adults, children), ambiguity type (PP, Adj), and the interaction of the two. The interaction was not significant (Est.=0.03, SE =0.25, z=0.12, p>.05). Because of this, two models were fitted with only group as the fixed effect and only ambiguity type as the fixed effect to calculate the main effects. The main effect of group was significant with adults scoring higher than children (Est.=-0.28, SE= 0.12, z=-2.31, p<.05). The main effect of ambiguity type was not significant (Est.= 0.11, SE=0.12, z=0.93, p>.05).

In sum, analyses revealed that adults significantly outperformed children in both total scores and scores on individual explanations. Participants did not perform significantly better on either ambiguity type in choosing the expected pictures or explanation scores, suggesting neither ambiguity type is easier to detect than the other. Finally, analyses show that there was no significant trend for children's overall scores to improve with age, but there was a significant trend for children's individual explanation scores to improve with age, indicating that children's explanations of syntactic ambiguity become more adult-like as they age.

4.0 General discussion

This study revealed two important findings. First, some 5- to 7-year-olds showed aptitude in ambiguity detection, which is earlier than previous research had found (e.g., Cairns et al., 2004). Nine out of the 55 children tested (~16%) gave adult-like explanations on at least half of the experimental items, scoring 6 or higher. Nine more children scored between 3 and 5, hinting at the beginning of development of syntactic ambiguity awareness. Another indication that some younger children are developing syntactic ambiguity awareness comes from their explanations of the adjective ambiguities. Four children, only one of whom was in the group of nine high-performers, exhibited the adult-like property of mentioning the underspecification of the adjective. Based on this, it seems likely that syntactic ambiguity detection begins to develop in different children at different times. In support of this, the data shows a developmental trend. Even though children were not more likely to choose both pictures on ambiguous items as they got older, there was a trend for them to produce more adult-like explanations of ambiguity as they got older. These results together suggest that some children develop syntactic ambiguity awareness before 8-years-old, and many more are beginning to develop it. This is younger than previous research on children's ambiguity detection predicts.

The second important finding from this study is the considerable variation in both children's and adults' awareness of syntactic ambiguity. Children's individual scores ranged from 0 to 9 (out of 12) with these spread across the whole range. Even among the 21 children who scored zero (i.e., who chose only one picture for ambiguous sentences), there was variation in how they approached the task. Some children gave detailed explanations describing why the target picture that they did not choose could not match the sentence. This suggests that those children had not accessed that interpretation of the sentence. Other children who seemed to adopt response strategies in which they chose one picture for each sentence and gave repeated undetailed responses (e.g., *because it matches*). Even the adult group showed variation with scores ranging from 3 to 12 (out of 12). Thus, my findings as a whole suggest that the development of syntactic ambiguity awareness is not monolithic. Its components develop for different children at different rates, and it is even varied among adults.

In sum, this study reveals that some 4- to 7-year-olds have syntactic ambiguity awareness and others are beginning to develop the ability during these years. Future research should address the specific factors that might influence children's syntactic ambiguity detection. For example, one limitation of this study is that it was conducted only in schools whose students tend to be from relatively high-income families. Previous research has shown that socioeconomic status influences both children's language and reading skills (e.g., Snow, Burns, & Griffin, 1998, 1998; National Center for Education Statistics, 1999; Dickinson & Tabors, 2001; Hoff, 2013; Calvo & Bialystok, 2014). Future research should probe the effect of socioeconomic status on children's syntactic ambiguity awareness in a more economically diverse population. Furthermore, future research should probe the connections between syntactic ambiguity detection and literacy development in children under 8-years-old. One promising avenue of research asks whether children under 8-years-old can be taught to detect syntactic ambiguity. Because metalinguistic awareness calls on conscious processes, it can be improved with training. While the current study included items for teaching the task, it did not involve any syntactic ambiguity detection training. Previous research on ambiguity detection training hints that even young children can be taught to notice ambiguity. For example, Yuill (2009) and Zipke et al. (2009) showed that 7- to 9-year-olds can be trained to better detect lexical and syntactic ambiguity using jokes and riddles that turn on ambiguities. Shakibai (2008) and Zipke (2011) showed that children as young as 5 and 6 could be trained to detect lexical ambiguities. Based on this, it seems plausible that 5- to 6-year-olds could learn syntactic ambiguity detection via training. If so, this may correspond to improved reading, as it did with the older children in Yuill (2009) and Zipke et al. (2009). This would be an important discovery as it would suggest certain practical applications. For example, syntactic ambiguity detection could be incorporated into early literacy curriculum just as other metalinguistic skills have (e.g., phonological and word awareness). Syntactic ambiguity detection could also be included in intervention strategies for young children who specifically struggle in sentence comprehension as opposed to decoding.

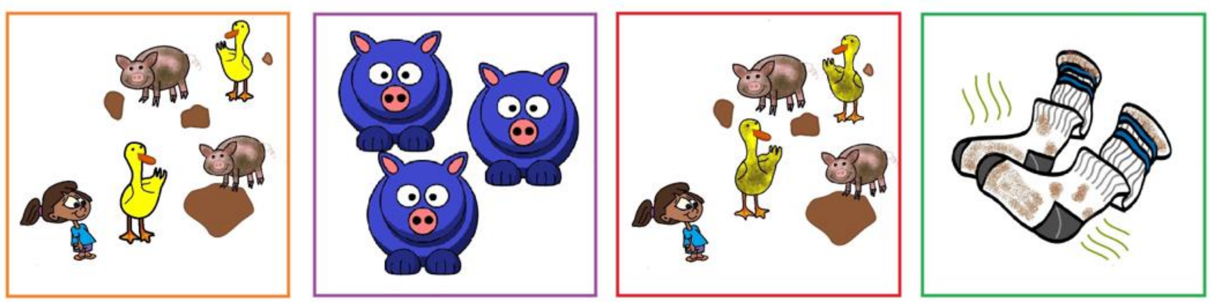
Appendix B1: Distractor images

The distractor images for all items were related to the sentence in some way. The specific properties of the distractors depended on the ambiguity type as shown in Table 4. Example pictures are shown in Figure 6 for an adjective ambiguity.

Table 4: Distractor images for ambiguous items.

	Ambiguous Adj Utterances (e.g., *The girl is noticing dirty pigs and ducks.*)	**Ambiguous PP Utterances** (e.g., *The man is pointing to the girl with the umbrella.*)
Dist. 1	first noun with different adjective (e.g., blue pigs)	same agent doing different action (e.g., man holding a flower)
Dist. 2	different noun with same adjective (e.g., dirty socks)	different agent with same instrument (e.g., grandma holding an umbrella)

Figure 6: Adjective ambiguity- *The girl is noticing dirty pigs and ducks.*

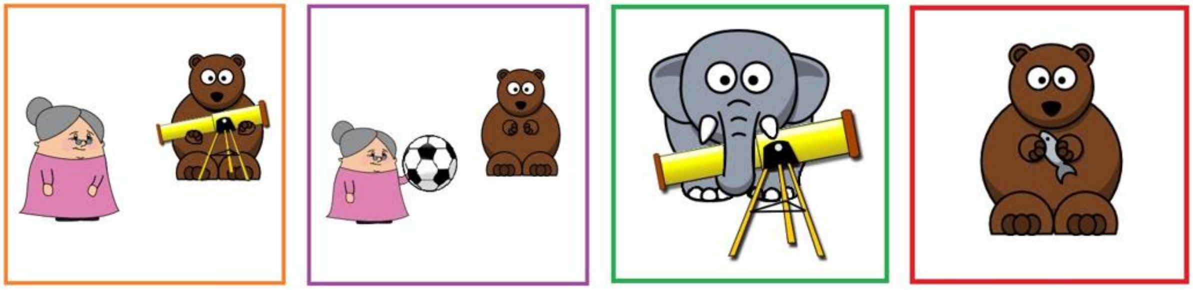


Again, for unambiguous items, all distractor images were related to the sentence in some way. One distractor image was more related to the sentence than the other two (highlighted in gray in Table 5). This was to offer a plausible second picture option on unambiguous items. An example of unambiguous PP item is shown in Figure 7.

Table 5: Distractor images for unambiguous items.

	Unambiguous Adj Utterances (e.g., *The chef is making strawberry cupcakes and strawberry ice cream.*)	**Unambiguous PP Utterances** (e.g., *The grandma is looking at the bear, and it has a telescope.*)
Dist. 1	second noun with different adjective (e.g., chocolate ice cream)	patient doing different action (e.g., bear holding a fish)
Dist. 2	different noun with same adjective (e.g., strawberry cookies)	different agent with same instrument (e.g., elephant holding telescope)
Dist. 3	same agent and action (e.g., chef making a cheeseburger & apple)	same agent and patient, different action (e.g., grandma giving bear a ball)

Figure 7: Unambiguous- *The grandma is looking at the bear, and it has a telescope.*



References

Aronsson, K. (1981). The bilingual preschooler as grammarian: Children's paraphrases of ungrammatical sentences. *Psychological Research Bulletin*, University of Lund.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7 [Computer Software]. Retrieved from http://CRAN.R-project.org/package=lme4

Bentin, S., Deutsch, A., & Liberman, I. Y. (1990). Syntactic competence and reading ability in children. *Journal of Experimental Child Psychology, 49*(1), 147-172.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development, 57*(2), 498-510.

Bialystok, E. (2001). Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics, 21*, 169-181.

Bialystok, E., & Ryan, E. B. (1985). Toward a definition of metalinguistic skill. *Merrill-Palmer Quarterly, 31*(3), 229-251.

Birdsong, D. (1989). *Metalinguistic performance and interlinguistic competence*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Cairns, H.S. (2013). Metalinguistic skills of children. In M. Becker, J. Grinstead, & J. Rothman (Eds.). *Generative Linguistics and Acquisition: Studies in Honor of Nina M. Hyams*, Philadelphia, PA: John Benjamins Publishing Company.

Cairns, H. S., Waltzman, D., & Schlisselberg, G. (2004). Detecting the ambiguity of sentences: Relationship to early reading skill. *Communication Disorders Quarterly, 25*, 68-78.

Cairns, H. S., Schlisselberg, G., Waltzman, D., & McDaniel, D. (2006). Development of a metalinguistic skill judging the grammaticality of sentences. *Communication Disorders Quarterly, 27*(4), 213-220.

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition, 130*(3), 278-288.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read?. *Cognition, 91*(1), 77-111.

Cazden, C. B. (1974). Play and metalinguistic awareness: One dimension of language experience. *The Urban Review, 7*(1), 28-39.

Demont, E., & Gombert, J. E. (1996). Phonological awareness as a predictor of recoding skills and syntactic awareness as a predictor of comprehension skills. *British Journal of Educational Psychology, 66*(3), 315-332.

Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school*. Baltimore, Md: P.H. Brookes Pub. Co.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*(3), 393-447.

Gerken, L., & Shady, M. E. (1996). The picture selection task. In D. McDaniel, C. McKee, & H.S. Cairns (Eds.). *Methods for assessing children's syntax*, Cambridge, MA: MIT Press.

Gleitman, L. R., & Gleitman, H. (1971). *Phrase and paraphrase: Some innovative uses of language* (1st ed.). New York, NY: Norton.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hove: Lawrence Erlbaum.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, *7*(1), 6–10.

Hakes, D. T. (1980). *The development of metalinguistic abilities in children*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Hermanto, N., Moreno, S., & Bialystok, E. (2012). Linguistic and metalinguistic outcomes of intense immersion education: how bilingual?. *International Journal of Bilingual Education and Bilingualism*, *15*(2), 131-145.

Hirsh-Pasek, K., Gleitman, L. R., & Gleitman, H. (1978). What did the brain say to the mind? Study of the detection and report of ambiguity by young children. In A. Sinclair, R. J. Jarvella, & W. J. M. Levelt (Eds.), *The child's conception of language*. New York: Springer Verlag.

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: implications for closing achievement gaps. *Developmental Psychology*, *49*(1), 4.

Hoover, W., & Gough, P. B. (1990). The Simple View of Reading. *Reading and Writing: An Interdisciplinary Journal*, *2*, 127–160.

Joshi, R. M., & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology*, *21*(2), 85-97.

Keil, F. (1980) Development of the ability to perceive ambiguities: evidence for the task specificity of a linguistic skill. *Journal of Psycholinguistic Research*, *9*, 219-30.

Kidd, E., Stewart, A., & Serratrice, L. (2011). Children do not overcome lexical biases where adults do: The role of the referential scene in garden-path recovery. *Journal of Child Language*, *38*, 222-234.

McDaniel, D., & Cairns, H. S. (1990). The child as informant: Eliciting linguistic intuitions from young children. *Journal of Psycholinguistic Research*, *19*, 331-334.

McDaniel, D., & Cairns, H. S. (1996). Eliciting judgments of grammaticality and reference. In D. McDaniel, C. McKee, & H.S. Cairns (Eds.). *Methods for assessing children's syntax*, Cambridge, MA: MIT Press.

Nation, K., & Snowling, M. J. (2000). Factors influencing syntactic awareness skills in normal readers and poor comprehenders. *Applied Psycholinguistics*, *21*(02), 229-241.

National Center for Education Statistics. (2011). *Achievement gaps: How Hispanic and white students in public schools perform in mathematics and reading on the national assessment of educational progress: Highlights*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education.

Omaki, A., & Lidz, J. (2015). Linking parser development to acquisition of syntactic knowledge. *Language Acquisition, 22*(2), 158-192.

R Core Team. (2014). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Schlisselberg, G. (1988). Development of selected conservation skills and the ability to judge sentential well-formedness in young children. (Unpublished Doctoral Dissertation). City University of New York, New York, NY.

Shakibai, M.T. (2008). The efficacy of a training program to teach kindergarteners lexical ambiguity detection. (Unpublished Doctoral Dissertation). City University of New York, New York, NY.

Shultz, T., & Pilon, R. (1973). Development of the ability to detect linguistic ambiguity. *Child Development, 44*, 728-733.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*, 238-299.

Snow, C., Burns, M.S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Thornton, R. (1996). Elicited production. In D. McDaniel, C. McKee, & H.S. Cairns (Eds.). *Methods for assessing children's syntax*, Cambridge, MA: MIT Press.

Tong, X., Deacon, S. H., & Cain, K. (2014). Morphological and syntactic awareness in poor comprehenders: Another piece of the puzzle. *Journal of Learning Disabilities, 47*(1), 22-33.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: studying online sentence processing in young children. *Cognition, 73*(2), 89-134.

Tunmer, W.E., Pratt, C., & Herriman, M.L. (1984). Metalinguistic abilities and beginning reading. *Reading Research Quarterly, 2*, 134-158.

van Kleeck, A. (1994). Metalinguistic development. In G. Wallach & K. Butler (Eds.). *Language learning disabilities in school-age children and adolescents: Some principles and applications.* New York: Macmillan.

Wankoff, L. S. (1983). Selected metalinguistic variables as they relate to conservation and reading achievement in normal achieving and some learning-disabled children. (Unpublished Doctoral Dissertation). City University of New York, New York, NY.

Wankoff, L. S., & Cairns, H. S. (2009). Why ambiguity detection is a predictor of early reading skill. *Communication Disorders Quarterly, 30*, 183-192.

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. Retrieved from http://arxiv.org/pdf/1308.5499.pdf.

Yuill, N. (2009). The relation between ambiguity understanding and metalinguistic discussion of joking riddles in good and poor comprehenders: Potential for intervention and possible processes of change. *First Language*, 29(1), 65-79.

Zimmer, E. (2016). Children's awareness of syntactic ambiguity. (Unpublished Doctoral Dissertation). University of Arizona, Tucson, AZ.

Zipke, M. (2011). First graders receive instruction in homonym detection and meaning articulation: The effect of explicit metalinguistic awareness practice on beginning readers. *Reading Psychology, 32*(4), 349-371.

Zipke, M., Ehri, L. C., & Cairns, H. S. (2009). Using semantic ambiguity instruction to improve third graders' metalinguistic awareness and reading comprehension: An experimental study. *Reading Research Quarterly*, *44*, 300-321.

Appendix C. Teaching First Graders to Detect Syntactic Ambiguity (Article 3)

Abstract

This six-week study investigates whether 6- to 7-year-olds can improve at syntactic ambiguity detection with practice, and whether such improvements correlate with improvements on reading readiness measures. Syntactic ambiguity detection correlates with reading comprehension and can be improved with practice in 8- to 9-year-olds. Although this skill might be a valuable addition to reading curricula and reading intervention measures, little is known about whether younger children have it or can develop it, or if the connection to reading begins before 8-years-old. Thus, the current study compares a group of first graders who did four weeks of syntactic ambiguity practice activities, to a group who did four weeks of math activities. Both groups took pre-/post-tests measuring syntactic ambiguity detection and reading readiness skills. The ambiguity group improved significantly more than the math group on both the syntactic ambiguity detection test and the reading readiness tests. This shows that children as young as 6 years of age can learn to detect syntactic ambiguity and suggests that syntactic ambiguity detection's connection to reading is in place then as well.

Keywords: syntactic ambiguity detection, children, first grade, language development, teaching, practice, reading readiness, literacy, metalinguistic skills

1.0 Introduction

Syntactic ambiguity occurs when an utterance has multiple meanings because it has multiple possible hierarchical structures. Consciously noticing such an ambiguity is called syntactic ambiguity detection. The development of syntactic ambiguity detection is an important topic because this skill likely supports children's reading development. Reading difficulties begin as early as pre-school (Snow, Burns, & Griffin, 1998; Torgesen, 2002; Lonigan & Phillips, 2016), and early reading success is a strong predictor of later reading success (Juel, 1988; Stanovich, 1986). One way to improve literacy curriculum and intervention strategies is to identify the many specific cognitive skills that contribute to reading success and how they interact with each other. For example, there are many metalinguistic precursors to literacy development (Goswami & Bryant, 1990; Bentin, Deutsch & Liberman, 1990; Ehri, Nunes, Stahl & Willows, 2001; National Reading Panel, 2000). The most well-known example is phonological awareness, or the ability to consciously identify and manipulate language sounds. It supports the decoding of letters to sounds of the language (Ehri et al., 2001; National Reading Panel, 2000). Because decoding is an important component of early reading development, phonological awareness has become an integral part of many reading curricula, most notably phonics-based programs (Baumann, Hoffman, Moon, & Duffy-Hester, 1998; see also Goswami & Bryant; Castles & Coltheart, 2004, for review).

Syntactic ambiguity awareness is another type of metalinguistic awareness that may contribute to reading. It was thought for many years that syntactic ambiguity detection did not emerge until as late as 12-years-old (Shultz & Pilon, 1973; Keil, 1980; Wankoff, 1983). In 2004, Cairns, Waltzman, & Schlisselberg lowered that age to second grade, or about 8-years-old. In my own work, I used a variation on a picture selection task to show that a small proportion of 4- to 7-year-olds can detect and describe syntactic ambiguity (Zimmer, 2016). In the task, children chose pictures that matched ambiguous or unambiguous utterances and then explained their choices to a puppet. I scored their explanations using a system based on properties of adult explanations. Specifically, adults in this task typically describe an explicit similarity and an explicit difference between their interpretations of ambiguous sentences and pictures. By explicit, I mean that adults referred to specific elements of the pictures that corresponded with the language in the sentence. About 16% of children in the study already could detect syntactic ambiguity, and others were showing indications of its development. For example, many children were able to select both pictures representing the two meanings of an ambiguity but their explanations exhibited only one or neither of the properties of adult explanations.

Both theoretical and empirical research suggest that, like phonological awareness, syntactic ambiguity detection supports reading development. The model of reading comprehension known as the Simple View of Reading (SVR) posits that two main components determine reading ability: decoding and linguistic comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990; Joshi & Aaron, 2000). Linguistic comprehension is the ability to interpret sentences and integrate them into the discourse. Helen Smith Cairns and colleagues treat the ability to detect structural ambiguity as part of linguistic comprehension (Cairns et al., 2004; Wankoff & Cairns, 2009; Cairns, 2013). Their reasoning is that if a child hears an ambiguous sentence and parses the wrong interpretation for the discourse context, it will reduce his or her overall listening comprehension. However, if a child can detect the ambiguity and change interpretations, his or her comprehension will be less affected. If this and the SVR are correct, then early structural ambiguity detection should correlate with successful reading comprehension.

Only a few empirical studies have probed connections between syntactic ambiguity detection and reading; they have all found positive correlations. Cairns et al. (2004) found that second graders' syntactic ambiguity detection scores correlated with their reading scores in the third grade. Yuill (2009) used jokes to train 8- to 9-year-olds to detect structural ambiguities. In her study, children were paired to play a computer game for several weeks; the game encouraged them to discuss jokes whose punchlines hinged on syntactic ambiguity. Each pair of children included a proficient reader and a struggling reader. She found that the children who improved at comprehending and discussing the ambiguity of the punchlines improved more at reading comprehension than their classmates who did not improve at noticing the ambiguity. Similarly, Zipke, Ehri, & Cairns (2009) trained third graders to detect syntactic ambiguity using jokes, riddles, picture discussion, and a passage-reading activity. They also found that improvements at syntactic ambiguity detection correlated with improvements on certain reading comprehension tests. Together, these studies suggest that syntactic ambiguity awareness does support reading comprehension, as Cairns and colleagues maintain.

Because of its connection to reading, we must better understand how this metalinguistic skill develops. If future research finds that there is a causal relationship between syntactic ambiguity awareness and reading improvement, then practice on ambiguity detection should be incorporated into reading curricula and reading intervention measures. In particular, practice on syntactic ambiguity awareness might be an intervention for children who struggle with reading comprehension but not with decoding (Cain, Oakhill, & Lemmon, 2005; Grigorenko, Klin, & Volmar, 2003; Stothard & Hulme, 1992; Kendeou, Savage, & Broek, 2009). However, there is still much we do not know about the development of syntactic ambiguity awareness and the specific ways that it connects to reading. For example, we do not know whether practice can affect children younger than 8-years-old. Yuill (2009) included a few 7-year-olds, but the mean age of participants was 9;1, and she does not report individual details about the 7-year-olds' improvement. One weak indication that younger children can learn the skill comes from Shakibai (2008) and Zipke (2011), who showed that children as young as 5 and 6 can learn to detect lexical ambiguity. It is critical that we investigate syntactic ambiguity detection in younger children because reading interventions are more effective the earlier they can be applied (Vellutino, Scanlon, & Tanzman, 1998). Thus, it is important that we know if children at earlier stages of reading can learn this potentially helpful skill.

The current study asks whether 6- to 7-year-olds' ambiguity detection can be improved with practice over a six-week period. To test for improvement, children were divided into an ambiguity group that played games eliciting discussion of syntactic ambiguity and a math group that played games eliciting discussion of mathematics. The ambiguity group was the test group, and the math group the control group. I measured improvement at syntactic ambiguity detection in both groups using a pre-/post-test that consisted of the task, materials, and scoring system from Zimmer (2016). The children also did pre-/post-tests on four markers of reading readiness to test whether improvement at syntactic ambiguity detection correlated with improvement on any other measures of emergent literacy (Hannah et al., 2009; Morris & EBooks, 2014). I compare how the scores of the children in the ambiguity group differ from those of the children in the math group. Table 1 outlines the schedule of the six-week long study, including the topics of each practice session.

Table 1: Schedule of study.

Week	**Ambiguity group**	**Math group**
1	Ambiguity pre-test Reading pre-tests	
2	Ambiguity practice 1: picture discussion	Math practice 1: counting
3	Ambiguity practice 2: sentence act out	Math practice 2: subitizing
4	Ambiguity practice 3: sentence diagram	Math practice 3: magnitude estimation
5	Ambiguity practice 4: family diagram	Math practice 4: addition and subtraction
6	Ambiguity post-test Reading post-tests	

I hypothesize that the ambiguity group will learn from the practice activities and thus improve on the ambiguity awareness test. I predict that the math group will not improve or will improve very little on the ambiguity awareness test. I also predict that the ambiguity group will improve more than the math group on the reading readiness tests. Whether or not the math group will improve on the reading readiness tests is an open question. On one hand, they may improve because they have an additional six weeks of schooling and reading experience. However, first graders are usually beyond receiving explicit instruction on reading readiness skills, so it is possible that the math group will not improve at all on the reading readiness tests. If only the ambiguity group improves on reading readiness, it seems reasonable to attribute that improvement to the ambiguity awareness practice.

2.0 Methods

I will now detail the study’s design, beginning with the participants and then moving on to the materials and procedures for the pre-/post-tests and the ambiguity and math practice activities.

2.1 Participants

The study included 40 6- to 7-year-olds: 21 participants in the ambiguity group and 19 in the math group. Twenty participants were originally in the math control group, but one was excluded because an audio recorder malfunction caused part of that child's data to be lost. Groups were balanced for gender, age, and bilingualism, but otherwise sorted randomly. The ambiguity practice group had 7 girls and 14 boys. The math control group had 6 girls and 13 boys. Both groups had a mean age of 6;8 at the end of the study. All participants were native speakers of English, though not all were monolingual. Each group included nine bilingual participants.[1] Participants were tested at private and charter elementary schools in Tucson, Arizona. Public school districts were not contacted because of time constraints for the research. It takes several months to gain permission to conduct research at schools in those districts. Private and charter schools only require the permission of the school directors.

2.2 Pre-test/post-test for syntactic ambiguity detection

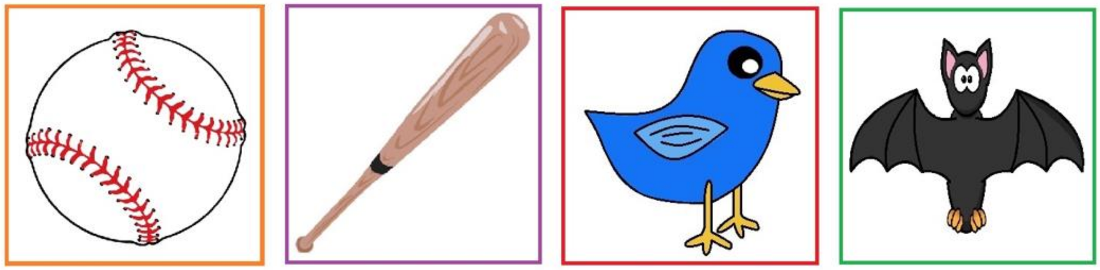
To measure participants' improvement in syntactic ambiguity detection, I used the variation on a picture selection task developed for Zimmer (2016). Two experimenters administered the test to one participant at a time. This test took about 15 minutes. One experimenter said sentences to the participants. For every sentence, participants saw four pictures on a board. They were instructed to choose all of the pictures that matched the experimenter's sentence. Two pictures matched each ambiguous item, corresponding to the two meanings of the ambiguous sentence. Only one picture matched each unambiguous item. So the target choices were always one or two pictures from the board. Experimenters checked with participants after they made their picture selection by asking if they were sure about their choice (e.g., *Are you sure that's the only picture that matches the sentence?*, *Are you sure there are two pictures that match the sentence?*). This checking was done regardless of what pictures the participant chose. After selecting pictures, participants explained their choices to a 'confused' puppet. We audio recorded and later transcribed and coded these explanations.

The task was introduced with a teaching dialogue that explicitly informed participants that language can be ambiguous, illustrating the point with a lexical example. Participants were then taught the task with three types of teaching items: a lexical ambiguity, a syntactically ambiguous sentence, and an unambiguous sentence. This teaching dialogue was intended to teach the task but not to provide significant training on ambiguity detection. The first teaching item of the game was the word *bat*. I taught the task with a lexical ambiguity item because previous research has shown that children are more aware of lexical ambiguity than syntactic ambiguity (Cairns et al., 2004; Wankoff & Cairns, 2009). Experimenters showed participants two target pictures representing the two meanings of *bat*, along with two distractor pictures, as in

[1] Eleven of the bilinguals were speakers of Spanish. The other bilinguals spoke French, German, Mandarin, Tamil, Hindi, and Arabic.

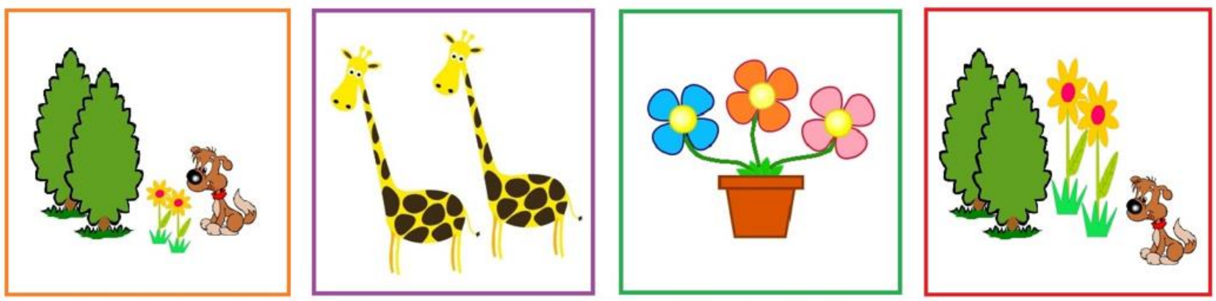
Figure 1. If the participant did not select both target pictures, experimenters gave the participant some feedback. Then, the sequence was repeated with a back-up lexical ambiguity item, *nail*.

Figure 1: Pictures for the lexical item *bat* reproduced from Zimmer (2016).



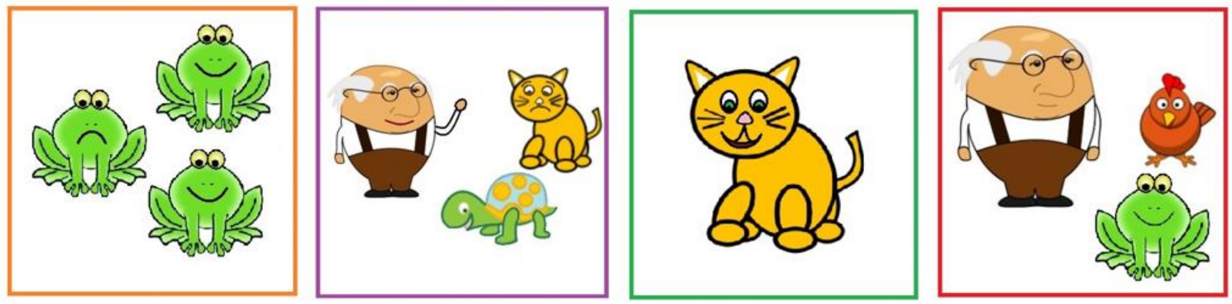
The next teaching item was an ambiguous sentence. Again, participants saw two target pictures representing the two interpretations of the sentence along with two distractors, as in Figure 2. If a participant failed to choose both pictures, experimenters gave some feed-back and did a back-up item (*The man is poking the monkey with banana*).

Figure 2: Pictures for ambiguous teaching item *The puppy is sniffing tall bushes and flowers* reproduced from Zimmer (2016).



The next teaching item was an unambiguous sentence. Participants saw one target picture that matched the sentence and three distractors, as shown in Figure 3. If a participant chose any picture besides the target, experimenters gave them some feedback and did a back-up item (*The boy is watching the girl, and she has a magnifying glass*). Once the teaching items were complete, participants continued to experimental items.

Figure 3: Pictures for unambiguous teaching item *The grandpa is waving at the happy turtle and the sad cat* reproduced from Zimmer (2016).



There were four ambiguous experimental sentences. These included two prepositional phrase attachment (PP) ambiguities (e.g., *The boy saw the girl with the binoculars*) and two pre-nominal adjective (Adj) ambiguities (e.g., *The teacher is studying green frogs and fish.*). There were also four unambiguous experimental sentences, as shown in Table 2.

Table 2: Experimental items for ambiguity awareness test.

Ambiguous Sentences		**Unambiguous Sentences**
PP	1. The man is pointing to the kid with an umbrella. 2. The boy is staring at the girl with binoculars.	5. The grandma is looking at the bear, and it has a telescope. 6. The woman is using a feather to tickle the baby.
Adj	3. The teacher is studying green frogs and fish. 4. The girl is noticing dirty pigs and ducks.	7. The doctor is wanting vanilla cakes and chocolate cookies. 8. The chef is making strawberry cupcakes and strawberry ice cream.

Recall that two pictures represented interpretations of each <u>ambiguous experimental sentence</u>. These were shown with two distractor images. One picture represented the meaning of each <u>unambiguous experimental sentence</u>. It was shown with three distractor images. For more about the task and materials, see Zimmer (2016).

For every ambiguous item, individual explanations were scored on a scale from 0-3, as shown below in Table 3. This scale was based on properties of adult explanations of ambiguities (Zimmer, 2016). Specifically, adults typically describe an explicit <u>difference</u> and explicit <u>similarity</u> between the two interpretations when explaining their picture choices in this task. Again, by explicit, I mean adult explanations refer to specific elements of the pictures that are relevant to the language of the sentence. For example, when describing a similarity for the sentence number 3 from Table 2, an adult would say “both pictures have green frogs” rather than a more generic similarity like “the pictures look the same.” Thus, explanations that referenced both an explicit similarity and an explicit difference received the highest possible score. All explanations were scored by two researchers who were blind to whether the participant was in the ambiguity group or the math group.

Table 3: Explanation scoring system, reproduced from Zimmer (2016).

Score	Picture Selection	Explanation
0	Selected only one target picture or any distractor picture.	N/A
1	Selected both and only the target pictures.	Participant discussed picture, but mentioned neither an explicit similarity nor an explicit difference.
2	Selected both and only the target pictures.	Participant discussed either an explicit difference or explicit similarity.
3	Selected both and only the target pictures.	Participant discussed both an explicit difference and explicit similarity.

To illustrate how this scoring system applies, Table 4 gives examples of children's explanations at every level of the scoring system (ages in parentheses). It also includes examples of adult explanations from Zimmer (2016) that scored a 2 or 3. No adult explanations from Zimmer (2016) scored a 1, so there is an empty cell.

Table 4: Examples of scored explanations.

Points	Examples of Child Explanations	Examples of Adult Explanations
1	"Because they match." (6;2)	N/A
2	"Both of them have fish and frogs." (6;7)	"Both of the pictures have frogs and fish."
3	"It can be the girl's noticing that there are dirty pigs and dirty ducks but it can also be clean ducks and dirty pigs." (6;11)	"Okay so the sentence could mean the girl is noticing dirty pigs and dirty ducks or it could mean the girl is noticing dirty pigs and just ducks."

This table illustrates how children's explanations of syntactic ambiguity can vary quite considerably in their similarity to adult responses. For the purpose of analyses, I calculated total scores for each participant at pre-test and post-test. A participant's total score was the sum of their scores on all four ambiguous items. With four experimental sentences, an individual's maximum total score was 12. In the results and analyses, I will report only total scores and not scores on individual items.

2.3 Pre-tests/post-tests for early literacy markers

To measure improvement in early literacy-related abilities, I used the following four tests of reading readiness: word detection, syllable detection, rhyming, and listening comprehension.

It is important to note that these are not tests of reading per se. Instead, they concern skills known to correlate with early reading success (Wagner & Torgesen; 1987; Goswami, 1990; 1999; Goswami & Bryant, 1990; Hoover & Gough, 1990; Sawyer, 1992; Wise, Sevcik, Morris, Lovett, & Wolf, 2007).

1. Syllable detection test: After listening to the researcher say a word, the participant segmented the word by tapping on a table for each syllable. This test is a measure of phonological awareness.
2. Rhyming test: After listening to the researcher say a word, the participant produced words that rhymed with the prompt. It tests whether participants can attach the same rime to different onsets. This test is a measure of phonological awareness.
3. Word detection test**:** After listening to the researcher read a sentence aloud, the participant repeated the sentence, clapping once on each word. Multisyllabic words were included to ensure participants could segment words as opposed to syllables. This test is a measure of lexical awareness.
4. Listening comprehension test: After listening to the researcher read a short story, the participant answered comprehension questions about what occurred. Participants could hear up to two stories. The second was more advanced than the first.

Two experimenters administered the reading readiness tests with one child at a time. This session took about 15 minutes. The instructions for all four tests were adapted from the Learning Disabilities Association of Alberta's *Reading Readiness Screening Tool* (Hannah et al., 2009). For the word detection, syllable detection, and rhyming tests, some items were used from *Reading Readiness Screening Tool.* I developed additional items to make the tests longer and increasingly difficult to detect a wider range in abilities. The stories and questions used in the listening comprehension test were both from the *Morris Informal Reading Inventory* (Morris & EBooks, 2014). Participants' responses on these tests were hand-recorded on paper and independently scored by both experimenters. Discrepancies between the scores were resolved using the audio recordings of the session. For more about the instructions and items for these measures, see Appendix A.

2.4 Practice materials for ambiguity group

As Table 1 summarizes, the ambiguity group participated in four practice sessions. Each session emphasized a different activity designed to encourage participants to discuss syntactic ambiguity. The sessions each involved two participants and two experimenters; one of the experimenters controlled a puppet. Each session lasted about 15 minutes, so the total practice time was about an hour distributed over four weeks.

Week 1: The first practice session for the ambiguity group involved the discussion of syntactic ambiguity using pictures. One experimenter said sentences to the participants. For each sentence, each participant received three pictures: one target picture and two distractors. Each participant selected the picture from his or her hand that matched the sentence. For ambiguous sentences, the two participants each had a different picture that matched the sentence, representing the two interpretations of the sentence. For unambiguous sentences, the two participants had the same picture that matched the sentence. Each participant then explained to the other why he or she picked that picture. Experimenters encouraged the participants to notice how two different pictures match an ambiguous sentence, but not an unambiguous sentence.

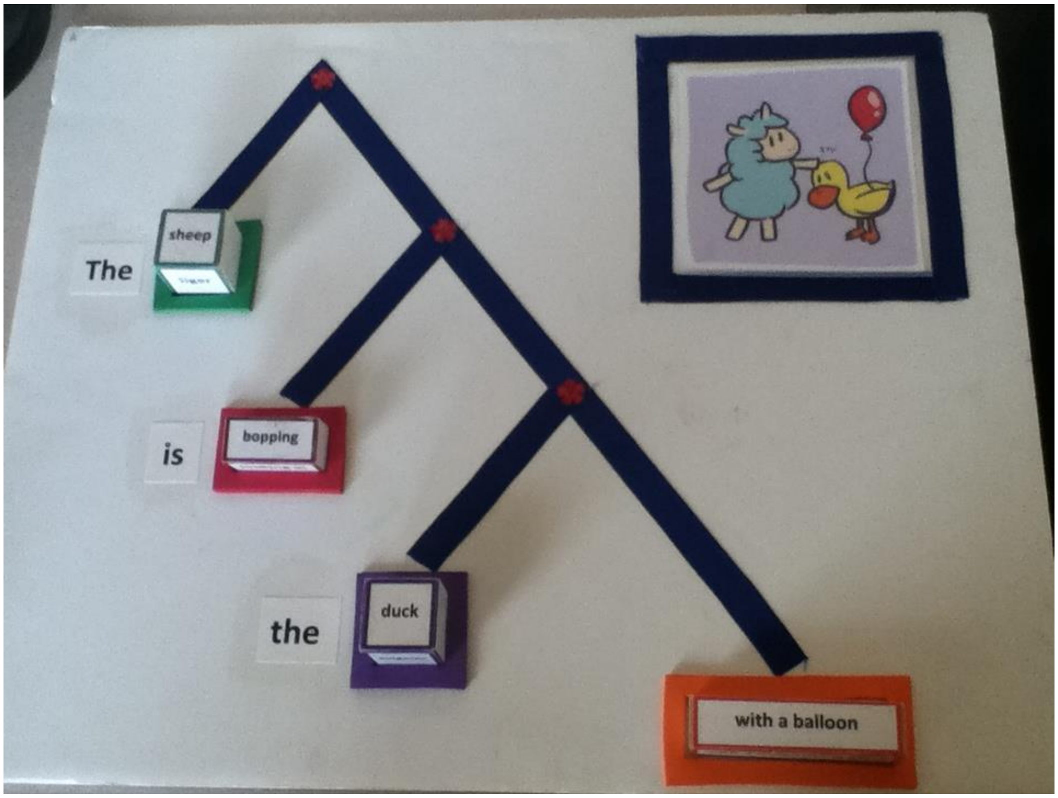
Week 2: The second practice session involved a game in which participants used toy props to act out sentences. Some sentences were syntactically ambiguous; others were unambiguous. Experimenters gave each participant a set of toys for every sentence. I manipulated the toys so that participants were encouraged to act out the sentence in different ways for ambiguous sentences and the same way for unambiguous sentences. For example, take the ambiguous sentence *The boy saw the lion with binoculars.* One participant received a set of toys that included a boy who had binoculars glued so that he was looking through them and a plain lion. Thus, this participant was likely to act out the interpretation of the sentence where the boy is using binoculars to see the lion. The other participant had a set of toys that included a plain boy and a lion that had binoculars glued around its neck. Thus, this participant was likely to act out the interpretation where the boy is looking at a lion that has binoculars. This is in contrast to the unambiguous sentence *The man is using a magnifying glass to look at the duck.* Each participant received an almost identical set of toys with a man, a magnifying glass, and a duck. Thus, children were likely to act out the sentence in basically the same way. Each participant then explained to the other why he or she acted out the sentence that way. Experimenters encouraged children to notice how they could act out an ambiguous sentence in two different ways.

Week 3: The third session involved physical representations of the syntactic structures of an ambiguous sentence. Each participant was placed in front of a different board with a very simplified syntax tree, representing one of the interpretations of a prepositional phrase attachment ambiguity.[2] The structures included no technical terms. The boards had slots in the place of the content words; function words were written directly on the boards. For each empty slot, there was a block with a different word or phrase on each side. Thus, the participant could create a sentence by choosing words with the blocks and putting them in the slots. Figure 4 is an image of one of the boards.

---

[2] This activity was based on one that my colleagues and I have designed for linguistics booths at festivals. In that activity, children match pictures representing the two interpretations of an ambiguous sentence to simplified tree structures. We find that children as young as 5-years-old can play the game and that they enjoy it too.

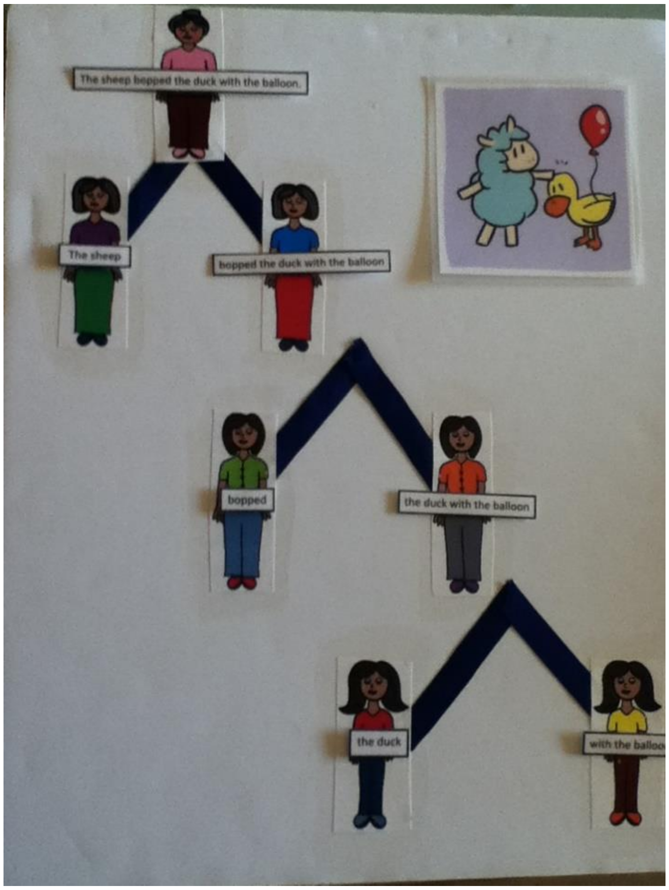
Figure 4: Modifier interpretation for *The sheep is bopping the duck with a balloon.*



In this game, experimenters gave each participant a picture representing one interpretation of an ambiguous sentence. The interpretation that each participant had depended on which board he or she had. The participants were then tasked with manipulating the blocks to create a sentence that matched the picture. Experimenters helped the small proportion of participants who had difficulty reading the blocks. After both participants found the right words, experimenters revealed that they had both created the same sentence for two different pictures. Experimenters then explained the trees to the participants to illustrate why the same sentence matched two different pictures. In this activity, participants learned to associate different interpretations with abstract structures.

Week 4: The final session also involved boards with syntactic structures for ambiguous sentences, but the representations of the sentence structure were based on familial relations this time. Thus each node of the board was represented by a member of a family, shown in Figure 5.

Figure 5: Modifier interpretation for *The sheep is bopping the duck with a balloon.*



I taught participants that some words in the structure are "sisters" because they come from the same "mother" in the sentence structure. After learning about the boards, participants matched the syntactic structures to the appropriate interpretations by selecting a picture.

2.5 Practice materials for math group

The math group also received about 15 minutes per practice session of interaction with the researchers. Like the ambiguity practice sessions, the math practice sessions each involved two participants and two experimenters; one of the experimenters controlled a puppet. Each session emphasized discussion of a different mathematics skill. Participants played the games in pairs to make them as similar as possible to the ambiguity group.

- **Week 1**: Participants practiced counting principles including counting forwards to 20 by one and by two and counting backwards from 10 to one.
- **Week 2**: Participants practiced subitizing (perceiving the size of a small set without counting).
- **Week 3**: Participants practiced judging the magnitude of numbers including estimating numbers on a blank number line and comparing set sizes (e.g., 2 is closer to 3 than 6).

- **Week 4**: Participants practiced combining and separating sets (i.e., addition and subtraction).

The four math skills I chose correlate with later success at mathematics (Jordan, Kaplan, Ramineni & Locuniak, 2009). I chose these mathematics skills for the control group because they require analytic reasoning as syntax does but are not directly related to literacy development (Purpura, Hume, Sims, & Lonigan, 2011). Some math skills, such as number word learning and story problems are correlated with literacy development (Purpura et al., 2011; Chang, Singh, & Filer, 2009). None of our activities focused on number word learning or story problems per se.

3.0 Results

In this section, I will emphasize comparisons of group scores but also mention individual variation. For the purpose of analysis, difference scores were calculated for each participant by subtracting the ambiguity and reading pre-test scores from the post-test scores. Positive difference scores show that the participant improved at the tests. The larger the difference score, the more improvement there was. Table 5 gives the mean pre-test, post-test, and difference scores for the ambiguity group and the math group on each test, including the ambiguity awareness test, each reading readiness test, and a composite score of all four reading readiness tests combined. The composite scores were calculated for the purpose of statistical analyses.

Table 5: Mean pre-test, post-test, and difference scores for all tests.

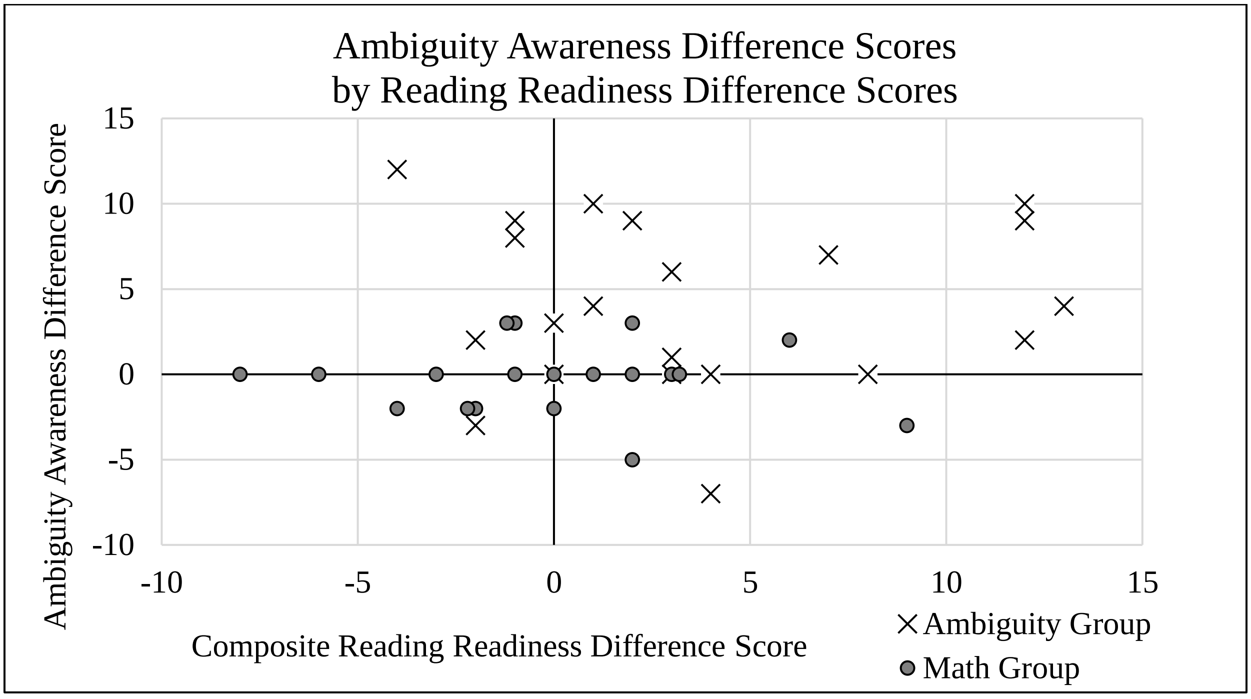
	Ambiguity Group (N=21)			**Math Group (N=19)**		
	Pre-test	Post-test	Diff.	Pre-test	Post-test	Diff.
Ambiguity awareness test	2.14 /12	6.24 /12	4.10	2.52 /12	2.26 /12	-0.26
Reading readiness (RR) tests						
Word detection	5.29 /9	6.10 /9	0.81	5.68 /9	5.68 /9	0.00
Syllable detection	5.24 /9	6.81 /9	1.57	7.05 /9	7.21 /9	0.16
Rhyming	7.10 /9	7.38 /9	0.29	6.63 /9	5.47 /9	-1.16
Listening comp.	5.90 /8	6.81 /8	0.90	5.79 /8	6.79 /8	1.00
Composite RR Score	23.5 /35	27.1 /35	3.57	25.2 /35	25.2 /35	0.00

To determine whether the differences between the ambiguity group and the math group were significant, Welch's t-tests comparing the two groups' difference scores were conducted in R (R Core Team, 2014). The first t-test compared the two groups' difference scores on the ambiguity awareness test. The ambiguity group had significantly higher difference scores than the math group ($t = 3.7$, $df = 27.5$, $p < .01$). In fact, the ambiguity group improved on the ambiguity awareness test by an average of 4.1 points. This is almost a 200% improvement from their average pre-test score of 2.1. This improvement is considerable in light of the fact that the math group did not show any improvement on the same test. A second t-test compared the two groups' composite reading readiness difference scores. The ambiguity group again had significantly higher difference scores than the math group (t=2.5, df= 37.1, p<.05). On average, this group improved by 3.6 points on their composite reading readiness score. Again, the math group did not improve at all on the composite reading score. We can thus conclude that participants learned from the ambiguity practice activities and this correlated with significant overall improvement on the reading readiness scores.[3] The suggests that participants in the ambiguity group improved as a result of the ambiguity awareness practice.[4]

I did not conduct statistics on the reading readiness tests individually because they each had only eight or nine items. However, the raw data shows that ambiguity group had higher difference scores on three of the four reading readiness tests. The only test on which the math group improved as much as the ambiguity group was the listening comprehension test, with the math group improving by 1 point and the ambiguity group improving by 0.9 points.

Figure 6 provides another way to see differences between groups; it also shows the individual variation among participants. It plots participants' ambiguity awareness difference scores (on the y axis) by their composite reading readiness difference scores (on the x axis).

Figure 6: Ambiguity awareness difference scores plotted by composite reading readiness difference scores and group.



[3] Wilcoxon rank sum tests were also used to compare the group difference scores on the ambiguity awareness test and the reading readiness tests. Unlike t-tests, these tests do not assume a normal distribution of the data. Results from both tests were also significant. The ambiguity group still had significantly higher difference scores than the math group for the ambiguity awareness tests (W=330.5, p<.01) and the reading readiness tests (W=278, p<.05).

[4] The study did not pre- or post-test the math skills due to time constraints. School directors requested that the study use under 30 minutes per week with each child.

From this plot, we can see that the math group is more centered around the intersection of the x-axis and y-axis than the ambiguity group is, which has points reaching up further into the upper right quadrant. All but four children in the ambiguity group improved at ambiguity awareness. Five children in the math group showed some improvement on the ambiguity awareness test. Both groups show considerable variation on the difference score for both the ambiguity awareness test and their composite reading readiness tests. In the ambiguity group, the difference scores for the ambiguity awareness test range from -7 to 12. This is a broad range given that the maximum possible difference score in either direction is 12. The difference scores for the reading readiness tests range from -4 to 13. In the math group, the difference scores for the ambiguity awareness test range from -5 to 3, and the difference scores for the reading readiness tests range from -8 to 9. Thus, even though the ambiguity group generally showed more improvement on both test types than the math group, individual children clearly varied in how effective the practice activities were for them.

Because research has shown that bilingualism correlates with advanced metalinguistic skills (e.g., Bialystok, 1986; 2001), I also compared the bilingual and the monolingual participants' ambiguity awareness pre-test scores. A Welch's t-test revealed no significant difference ($t = -1.57$, $df = 31.56$, $p>0.05$), although the mean for bilinguals was higher at 3.1 than for monolinguals at 1.7. The number of participants in each group (ambiguity and math) was too small to do statistical analyses comparing monolinguals to bilinguals within the groups separately. The raw numbers suggest that monolinguals in the ambiguity group may have improved more than bilinguals in the ambiguity group did. Monolinguals in the ambiguity group had a mean difference score of 4.9 on the ambiguity awareness test; bilinguals in the same group had a mean difference score of 3. The monolinguals in the ambiguity group also improved more on the reading readiness tests with a mean difference score of 4 compared to bilinguals in the same group with a mean difference score of 3. This hints that monolinguals may have benefited more from the training than their bilingual counterparts. However, because the number of bilingual participants in each group was too small to conduct comparative statistical analyses, I cannot say if these differences were significant.

Before turning to a general discussion of these results, it is important to address how the pre-test scores differed across the two groups. The groups had similar ambiguity awareness pre-test scores, at 2.1 for the ambiguity group and 2.5 for the math group. This difference was not significant based on a Welch's t-test ($t=-0.45$, $df = 37.9$, $p>0.5$). Although the raw numbers suggest a slight difference in the composite reading readiness scores of the two groups (ambiguity group at 23.5 and math group at 25.2), this difference was also not significant based on a Welch's t-test ($t = -0.78$, $df = 38.9$, $p>.05$). Thus, each group's improvement does not reflect differences on their pre-test performances.[5]

4.0 General Discussion

This study shows that 6- to 7-year-olds can improve at syntactic ambiguity awareness with training. From only one hour of syntactic ambiguity detection practice, the ambiguity group improved on the ambiguity awareness test an average of 4.1 points on a 12-point scale; the math group did not improve. Improvement at syntactic ambiguity awareness also correlates with improvement on reading readiness tests. The ambiguity training group improved an average of 3.6 points on a 35-point scale. Again, the math control group did not improve. Because the only systematic differences between these two groups were their practice activities, this suggests that children did indeed learn from the syntactic ambiguity awareness practice. These results are important for two reasons. First, this lowers the age that we know children can learn to detect syntactic ambiguity. Previous research showed that 8- to 9-year-olds could improve at the skills with practice (Yuill, 2009; Zipke et al., 2009). Yuill (2009) included a few 7-year-olds, but the training group in the study had a mean age of 9;1. The mean age of the children in the current study was 6;8. This age distinction matters because previous research has shown that many children are beginning to develop spontaneous syntactic ambiguity detection naturally (as opposed to via training) at about 8-years-old (Cairns et al., 2004). Thus, the current study shows that children can be taught ambiguity detection before the ability shows itself naturally.

The findings from the reading readiness tests support previous research that suggests that syntactic ambiguity awareness contributes to reading development (e.g., Cairns et al., 2004; Yuill, 2009). In the current study, children who practiced syntactic ambiguity awareness improved significantly on the reading readiness tests, which measured known markers of emergent literacy. Interestingly, the three reading readiness tests that the ambiguity group improved on relied on metalinguistic skills: word detection, syllable detection, and rhyming. This suggests that improvement at syntactic ambiguity detection supports the development of other metalinguistic skills. Such reading readiness tests, however, are a somewhat indirect measure of how syntactic ambiguity awareness contributes to reading. A more direct measure would be standardized reading tests. This would likely require a longer study than the current six-week study as such tests are generally administered once per year.

Both groups improved about the same amount on the listening comprehension test. At face value, this seems inconsistent with the hypothesis that better syntactic ambiguity awareness should improve children's listening comprehension (e.g., Cairns et al., 2004). This hypothesis

[5] I also calculated mean ambiguity awareness difference scores for boys and girls separately to determine if gender affected improvement on the ambiguity awareness test. In the ambiguity group, the mean ambiguity awareness difference score was 4.14 for girls and 4.07 for boys. In the math group, the mean ambiguity awareness difference score was -0.17 for girls and -0.31 for boys. This comparison suggests that the practice activities did not work better for either gender.

was grounded in the Simple View of Reading. Recall that this view posits that listening comprehension is one of the two main components of reading. Cairns and colleagues theorize that syntactic ambiguity detection helps people overcome comprehension difficulties caused by ambiguity and thus should improve reading comprehension. In this experiment, I did not directly test the Simple View of Reading's premise that listening comprehension is a main component of reading. I intended to test Cairns et al.'s hypothesis that syntactic ambiguity awareness contributes to reading by showing that it leads to improvement on skills that correlate with reading success. However, the listening comprehension test that I chose did not directly test Cairns et al.'s more specific hypothesis that syntactic ambiguity awareness helps children overcome listening comprehension difficulties that are caused by ambiguity. The listening comprehension test intentionally did not include syntactically ambiguous sentences. This is because I wanted the test to be a measure of listening comprehension independent of syntactic ambiguity awareness. To better test Cairns et al.'s specific hypothesis, future research should probe the effect of ambiguity awareness on children's listening comprehension of stories that include ambiguities.

One possible criticism of my interpretation of the ambiguity groups' improvement on the reading readiness tests is that it was not due to improvement on (syntactic) ambiguity detection per se. Rather, it might have been due to the verbal nature of the ambiguity activities. In other words, the ambiguity groups' improvement on reading readiness measure may have been due to a general improvement in verbal skills. Future research should probe this distinction. One way to do so would be to conduct a study similar to this one with a control group that does activities with a focus on verbal skills that are not directly related to syntactic ambiguity detection, such as vocabulary development, poetry, and story-telling. Like in the current study, the test group would do activities to practice syntactic ambiguity detection per se. If the test group that practiced ambiguity detection improved more on the reading readiness measures than the control group who practiced other verbal skills, it would indicate that syntactic ambiguity awareness contributes independently to reading development.

In sum, this work suggests that syntactic ambiguity detection is teachable and thus could be a part of reading curricula or intervention strategies if future research clarifies how it contributes to reading comprehension. More research should address the relationship of syntactic ambiguity awareness and reading. Since reading intervention strategies are more effective when applied before the first grade, it would be valuable to know if children in preschool or kindergarten could learn to detect syntactic ambiguity. Future research should examine whether this skill could be taught even younger. Another line of future research might examine how other factors interact with syntactic ambiguity detection. For example, a limitation of this study is that it was conducted in schools with relatively homogenous student populations in terms of socioeconomic status, which was high compared to Tucson-area public school districts. Since socioeconomic factors are known to affect language skills and reading development (Snow et al., 1998; National Center for Education Statistics, 1999; Dickinson & Tabors, 2001; Hoff, 2013; Calvo & Bialystok, 2014), future research should take place at schools with economically more diverse students. Another factor that effects development of metalinguistic skills is multilingualism. Many multilinguals tend to have better metalinguistic skills than monolinguals (Aronsson, 1981; Bialystok, 1986; 2001; Davidson, Raschke, & Pervez, 2010; Hermanto, Moreno, & Bialystok, 2012). Although not all multilinguals benefit evenly in their metalinguistic skills. For example, child brokers, who translate for their parents have particularly advanced metalinguistic abilities compared to other multilinguals (Raschke, 2013). Balanced bilinguals have stronger metalinguistic skills than unbalanced bilinguals (Bialystok & Barac, 2012). Similarly, bilinguals who are literate in both languages have better metalinguistic abilities than bilinguals who are literate in only one (Thomas, 1988). Thus, future research should probe how different types of multilingualism influence children's natural syntactic ambiguity awareness and its contribution to reading development. Other cognitive factors that could influence ambiguity awareness are working memory and general intelligence. By investigating the specific components that contribute to syntactic ambiguity detection, we can better understand how to incorporate it into reading curricula for diverse children.

Appendix C1: Reading readiness test materials

The instructions for all four reading readiness tests were adapted from the Learning Disabilities Association of Alberta's *Reading Readiness Screening Tool* (Hannah et al., 2009, see www.righttoread.ca). For the word detection, syllable detection, and rhyming tests, some items were used from *Reading Readiness Screening Tool.*[6]

1. Word detection test: The test included 9 items. The first four items of the test were from Hannah et al. (2009). I developed the rest. We stopped the test when a participant missed two in a row.

Instructions: *I am going to say a sentence. I want you to clap one time for each word I say in the sentence. We'll try together first and then you try on your own.*
Here is the sentence: The book is green.
OK, are you ready to clap with me? (Child and researcher clap together) *Now you try one.*

Table 6: Examples of word detection test items.

Sentence	Target
The gorilla is eating a banana.	6 claps
The monkey is in the big tree.	7 claps
We like to play with the purple toy.	8 claps

2. Syllable detection test items: The test included 9 items. The first four items of the test were from Hannah et al. (2009). I developed the rest. We stopped the test when a participant missed two in a row.

Instructions: *I'm going to say a word. I want you to tap on the table for each part of the word I say. We'll try together first and then you try on your own. Here is the word:* November.
Are you ready to tap with me? (Child and researcher tap together) *Now you try one.*

Table 7: Examples of syllable detection test items.

Word	Target
Cat	1 taps
Rhinoceros	4 taps
Elephant	3 taps
Rainbow	2 tap

[6] I received permission to include these instructions from the coordinator of the Right to Read Program for the Learning Disabilities Association of Alberta, Canada (Heckbert, personal communication, 8 April 2016).

3. Rhyming test: The test included 9 items. The first five items of the test were from Hannah et al. (2009). I developed the rest. We stopped the test when a participant missed two in a row.

Instructions: *I'm going to say a word. I want you to say a wor*d *that rhymes with it. You can make up a word if you want.*

Table 8: Examples of rhyming test items.

Word	**Target**
fast	*last, past, cast, etc.*
swing	*thing, sing, etc.*
bird	*word, third, etc.*
eagle	*beagle, seagull, etc.*

4. Listening comprehension test: After listening to the researcher read a short story, the child answered comprehension questions about what occurred. Participants could hear up to two stories. Participants only heard the second story if they answered correctly on three out of four questions about the first story. The stories and questions used in the listening comprehension test were both from the *Morris Informal Reading Inventory* (Morris & EBooks, 2014). The first story was *Clouds* on page 189. The second and more advanced story was *Sledding* on page 193.

References

Aronsson, K. (1981). The bilingual preschooler as grammarian: Children's paraphrases of ungrammatical sentences. *Psychological Research Bulletin*, University of Lund.

Baumann, J.F., Hoffman, J.V., Moon, J., & Duffy-Hester, A.M. (1998). Where are teachers' voices in the phonics/whole language debate? Results from a survey of U.S. elementary classroom teachers. *The Reading Teacher, 51*, 636-650.

Bentin, S., Deutsch, A., & Liberman, I. Y. (1990). Syntactic competence and reading ability in children. *Journal of Experimental Child Psychology, 49*(1), 147-172.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development*, *57*(2), 498-510.

Bialystok, E. (2001). Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics*, *21*, 169-181.

Bialystok, E., & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition*, *122*(1), 67-73.

Cain, K., Oakhill, J., & Lemmon, K. (2005). The relation between children's reading comprehension level and their comprehension of idioms. *Journal of Experimental Child Psychology*, *90*(1), 65-87.

Cairns, H.S. (2013). Metalinguistic skills of children. In M. Becker, J. Grinstead, & J. Rothman (Eds.). *Generative Linguistics and Acquisition: Studies in Honor of Nina M. Hyams*, Philadelphia, PA: John Benjamins Publishing Company.

Cairns, H. S., Waltzman, D., & Schlisselberg, G. (2004). Detecting the ambiguity of sentences: Relationship to early reading skill. *Communication Disorders Quarterly, 25*, 68-78.

Cairns, H. S., Schlisselberg, G., Waltzman, D., & McDaniel, D. (2006). Development of a metalinguistic skill judging the grammaticality of sentences. *Communication Disorders Quarterly, 27*(4), 213-220.

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, *130*(3), 278-288.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition, 91*(1), 77-111.

Chang, M., Singh, K., & Filer, F. (2009). Language factors associated with achievement grouping in math classrooms: A cross-sectional and longitudinal study. *School Effectiveness and School Improvement*, *20*, 27–45.

Davidson, D., Raschke, V. R., & Pervez, J. (2010). Syntactic awareness in young monolingual and bilingual (Urdu–English) children. *Cognitive Development*, *25*(2), 166-182.

Demont, E., & Gombert, J. E. (1996). Phonological awareness as a predictor of recoding skills and syntactic awareness as a predictor of comprehension skills. *British Journal of Educational Psychology, 66*(3), 315-332.

Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school*. Baltimore, Md: P.H. Brookes Pub. Co.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*(3), 393-447.

Goswami, U. (1990). A special link between rhyming skill and the use of orthographic analogies by beginning readers. *Journal of Child Psychology and Psychiatry*, *31*(2), 301-311.

Goswami, U. (1999). Causal connections in beginning reading: The importance of rhyme. *Journal of Research in Reading*, *22*(3), 217-240.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read.* Hove: Lawrence Erlbaum.

Grigorenko, E. L., Klin, A., & Volkmar, F. (2003). Hyperlexia: disability or superability?. *Journal of Child Psychology and Psychiatry*, *44*(8), 1079-1091.

Hannah, S., Craig, J., Heckbert, J., Pentyliuk, M., Burke, K., & Siegel, L. (2009). *Reading Readiness Screening Tool*. Alberta, CA: Learning Disabilities Association of Alberta.

Hermanto, N., Moreno, S., & Bialystok, E. (2012). Linguistic and metalinguistic outcomes of intense immersion education: how bilingual? *International Journal of Bilingual Education and Bilingualism*, *15*(2), 131-145.

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: implications for closing achievement gaps. *Developmental Psychology*, *49*(1), 4.

Hoover, W., & Gough, P. B. (1990). The Simple View of Reading. *Reading and Writing: An Interdisciplinary Journal, 2*, 127–160.

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*(3), 850-867.

Joshi, R. M., & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology, 21*(2), 85-97.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, *80*(4), 437.

Keil, F. (1980) Development of the ability to perceive ambiguities: evidence for the task specificity of a linguistic skill. *Journal of Psycholinguistic Research*, *9*, 219-30.

Kendeou, P., Savage, R., & Broek, P. (2009). Revisiting the simple view of reading. *British Journal of Educational Psychology*, *79*(2), 353-370.

Lonigan, C. J., & Phillips, B. M. (2016). Response to instruction in preschool: Results of two randomized studies with children at significant risk of reading difficulties. *Journal of Educational Psychology*, *108*(1), 114.

Morris, D., & Ebooks Corporation. (2014). *Morris informal reading inventory: Preprimer through grade 8*. New York, NY: The Guilford Press.

National Center for Education Statistics. (2011). *Achievement gaps: How Hispanic and white students in public schools perform in mathematics and reading on the national* *assessment of educational progress: Highlights*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education.

National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.

Purpura, D. J., Hume, L. E., Sims, D. M., & Lonigan, C. J. (2011). Early literacy and early numeracy: The value of including early literacy skills in the prediction of numeracy development. *Journal of experimental child psychology*, *110*(4), 647-658.

R Core Team. (2014). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Sawyer, D. J. (1992). Language abilities, reading acquisition, and developmental dyslexia: A discussion of hypothetical and observed relationships. *Journal of Learning Disabilities*, *25*(2), 82-95.

Shakibai, M.T. (2008). The efficacy of a training program to teach kindergarteners lexical ambiguity detection. (Unpublished Doctoral Dissertation). City University of New York, New York, NY.

Shultz, T., & Pilon, R. (1973). Development of the ability to detect linguistic ambiguity. *Child Development*, *44*, 728-733.

Snow, C., Burns, M.S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360-407.

Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children. *Reading and Writing*, *4*(3), 245-256.

Thomas, J. (1988). The role played by metalinguistic awareness in second and third language learning. *Journal of Multilingual & Multicultural Development*, *9*(3), 235-246.

Thornton, R. (1996). Elicited production. In D. McDaniel, C. McKee, & H.S. Cairns (Eds.). *Methods for assessing children's syntax*, Cambridge, MA: MIT Press.

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of school psychology*, *40*(1), 7-26.

Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1998). The case for early intervention in diagnosing specific reading disability. *Journal of School Psychology*, *36*(4), 367-397.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*(2), 192.

Wankoff, L. S. (1983). Selected metalinguistic variables as they relate to conservation and reading achievement in normal achieving and some learning-disabled children. (Unpublished Doctoral Dissertation). City University of New York, New York, NY.

Wankoff, L. S., & Cairns, H. S. (2009). Why ambiguity detection is a predictor of early reading skill. *Communication Disorders Quarterly, 30*, 183-192.

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., & Wolf, M. (2007). The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *Journal of Speech, Language, and Hearing Research, 50*(4), 1093-1109.

Yuill, N. (2009). The relation between ambiguity understanding and metalinguistic discussion of joking riddles in good and poor comprehenders: Potential for intervention and possible processes of change. *First Language*, 29(1), 65-79.

Zimmer, E. (2016). Children's awareness of syntactic ambiguity. (Unpublished Doctoral Dissertation). University of Arizona, Tucson, AZ.

Zipke, M. (2011). First graders receive instruction in homonym detection and meaning articulation: The effect of explicit metalinguistic awareness practice on beginning readers. *Reading Psychology, 32*(4), 349-371.

Zipke, M., Ehri, L. C., & Cairns, H. S. (2009). Using semantic ambiguity instruction to improve third graders' metalinguistic awareness and reading comprehension: An experimental study. *Reading Research Quarterly, 44*, 300-321.