Children take only some sentences literally: Investigating children's variable performance with scalar inferences

Cory Bill

A thesis submitted in fulfilment
of the requirements for the degree of
Doctor of Philosophy

Department of Cognitive Science
Macquarie University
Australia

August 2017

Contents

Preliminary content

Summary

This thesis investigates young English-speaking children's understanding of several kinds of *scalar inferences*. For adult speakers, scalar inferences are derived in response to sentences that have been introduced in a conversational context by negating alternative sentences that are more informative, but were not asserted. The findings of previous studies have revealed significant variation in children's sensitivity to scalar inferences. A family of explanations called the *Alternatives-based approach* attribute children's variable performance in computing scalar inferences to differences in the ease with which they can compose the relevant alternatives. A basic claim of this approach is that children more readily compute scalar inferences when they can assemble the alternatives using elements from the asserted sentence. The experiments reported in this thesis were designed to evaluate this claim by investigating children's performance with a series of untested inferences. The inferences that are tested in the thesis include *distributive* and *conjunctive inferences* (Chapter 2), two kinds of *free choice inferences* (Chapter 3), and two inferences associated with sentences where 'some' is in the scope of 'every' (Chapter 4). The results of these investigations support the Alternatives-based approach by demonstrating that children readily compute inferences in which the alternatives can be formulated using parts of the asserted sentence. The results also provide contributions to several other important theoretical issues.

Declaration

I, Cory Bill, for the work presented in this thesis, can confirm that:

- This work has not been submitted for a degree or any other qualification at this university or any other institution.
- Where I have consulted the work of others, this is always clearly attributed.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.
- Ethics approval has been obtained from The Faculty of Human Sciences Human Research Ethics Sub-Committee (Ref: 5201300073) for the research presented.

Signed: CBill

Date: 16/08/17

Acknowledgments

First, I need to thank my primary supervisor, Stephen Crain. Stephen has been a wonderful support, and a great mentor. He has taught me so much about linguistic theory, how to conduct child language acquisition research, as well as the application of general scientific principles. I especially want to thank Stephen for the support he gave me in my final year when my progress was interrupted by a series of unexpected and unfortunate medical emergencies in my family.

I also need to thank my co-supervisor Rosalind Thornton. Rozz started me on this path when I was still doing my undergraduate degree. It was in her language acquisition course that I was first introduced to research on child language acquisition and generative linguistic theory. In fact, it was during this period that she got me involved in my first experiment investigating child language. Rozz has been a great support, particularly when I needed help making some of my experimental items more interesting.

During the first year of my PhD, Jacopo Romoli was employed as a post-doc at Macquarie University and was one of my co-supervisors. Jacopo helped me immensely during those early days, as I was first becoming acquainted with the formal semantics/pragmatics literature. I still miss being able to just walk over to his office and ask his help interpreting a section of a paper that I was struggling with. Jacopo has also been a great support for me as I entered and began navigating the academic world, helping me write my first conference and journal article submission, and inviting me to be a part of numerous collaborations.

Elena Pagliarini also deserve a great deal of thanks. She is a very talented researcher, and I have learnt a lot from her. I am particularly impressed by her sound understanding of scientific principles and how to apply them to linguistic research. I especially need to thank Elena for her contributions to the overall design of the experiments and to the creation of the experimental materials presented in this thesis.

Another person I need to thank is Lyn Tieu. Lyn was just finishing her PhD as I was beginning mine. She is a great researcher, and has helped to some extent with all of the research presented in this thesis. I especially want to thank Lyn for her help with the statistical analyses presented in this thesis.

I also want to thank the many people who given me feedback on this research, and have helped me develop the skills that made it possible, including: Florian Schwarz, Uli Sauerland, Teresa Guasti, Nicole Gotzner, Alexandre Cremers, and Jeremy Zehr. They are all amazing researchers, mentors, and friends. I have learnt so much from them, and they all helped to make this thesis possible.

The Language Acquisition Lab as Macquarie University is a wonderful group of researchers. It was a great space where I could pitch and get feedback on new experiment ideas or on interpretations of recently collected data. I want to thank all the members of this lab including: Anna Notely, Peng Zhou, Vincenzo Moscati, Aijun Huang, Likan Zhan, Mike Iverson, Shasha An, David (Haiquan) Huang, Na Gao, Loes Koring, and anyone else that I have missed. Their feedback, support and friendship has been invaluable.

There are also a number of other friends that I need to thank, including; Ben, Michael, Owen, Bernice, and Caitlin. Their constant emotional support and encouragement has helped me stay (relatively) sane throughout this journey.

I want to also acknowledge the support given to me by my mum and dad, my brother, Shanon, my sisters, Tiarn and Hannah, and my grandparents. Even though they did not always understand quite what I spend so much of my time doing, I am grateful for their love and support, particularly in this last year, when my life has become consumed by finishing this project. I also want to thank my parents-in-law, Grahame and Nenette, for all their support and encouragement.

Finally, I need to thank my wife, Giselle. In one sense, she bears the most responsibility for my current situation. A little over 10 years ago, not long after we had started dating, I was deciding what to do after graduating high school. Giselle encouraged me to think seriously about going to university, and assured me that I had the ability to succeed there. Giselle's constant love, understanding, and encouragement has both driven me to keep pushing forward, as well as given me the capacity to do so. Our lives together have had many different chapters. This most recent chapter, where we both decided to do PhDs simultaneously, has certainly not been the easiest. However, it has re-affirmed to me how incredible she is, and how much I love having her in my life. I look forward to our next chapter.

List of original papers

A number of the chapters in the present thesis are based on papers I created with the help of co-authors. I have identified below each of these papers, as well as the thesis chapter they are associated with. Also, for each of these papers, I have described my contribution, as well as the contribution of my co-authors to their creation. A summary is also presented in Table 1. Authors are referred to by their initials.

Chapter 2 (I): Pagliarini, E., **Bill, C.**, Romoli, J., Tieu, L., and Crain, S. (2017). On children's variable success with scalar inferences: Insights from disjunction in the scope of a universal quantifier. Under review.
Author contributions:
All authors assisted in the conception and design of the experiment. EP and **CB** created the materials for the experiment, with feedback from SC. **CB** conducted the data collection, with assistance from EP. EP conducted the statistical analysis, assisted by **CB** and LT. EP drafted the results section of the paper, all other sections of the paper were drafted by **CB**. Paper revisions were contributed to by all authors by providing feedback, as well as by directly editing the paper.

Chapter 3 (II): Bill, C., Pagliarini, E., Romoli, J., Tieu, L., and Crain, S. (2017). Children's derivation of universal free choice inferences. In preparation.
Author contributions:
All authors assisted in the conception and design of the experiment. **CB** and EP created the materials for the experiment, with feedback from SC. **CB** conducted the data collection, with assistance from EP. **CB** conducted the statistical analysis, assisted by LT. All sections in the paper were drafted by **CB**. Paper revisions were contributed to by all authors by providing feedback, as well as by directly editing the paper.

Chapter 4 (III): Bill, C., Pagliarini, E., Romoli, J., Tieu, L., and Crain, S. (2017). Children's interpretation of sentences with multiple scalar terms. In preparation.

Author contributions:
All authors assisted in the conception and design of the experiment. **CB** and EP created the materials for the experiment, with feedback from SC. **CB** conducted the data collection, with assistance from EP. **CB** conducted the statistical analysis, assisted by LT. All sections in the paper were drafted by **CB**. Paper revisions were contributed to by all authors by providing feedback, as well as by directly editing the paper.

	I	II	III
Conception & design	EP, **CB**, JR, SC, LT	**CB**, EP, JR, SC, LT	**CB**, EP, JR, SC, LT
Planning & implementation	**CB**, EP	**CB**, EP	**CB**, EP
Data collection	**CB**, EP	**CB**	**CB**
Analysis & interpretation	EP, **CB**, LT	**CB**, LT, EP	**CB**, LT
Writing the paper	**CB**, EP, LT, SC, JR	**CB**, EP, SC, LT, JR	**CB**, EP, SC, LT, JR
Overall responsibility	EP, **CB**	**CB**	**CB**

Table 1: Division of labour in co-authored articles. Papers are labelled using the roman numerals attached to them in the list above.

Chapter 1

Introduction

1.1 Introduction

In everyday conversations, language users often derive inferences that extend beyond the literal content of the assertions made by the interlocutors. For example, the assertion of a declarative sentence with disjunction, such as (1), is true if just one, or if both, of the disjuncts are true. However, when language users interpret the assertion in (1), they are likely to include the inference that only one of the disjuncts are true, as indicated in (2). This *exclusivity inference* is a classic example of a group of inferences known as *scalar inferences*.[1]

(1) Mary brought beer or wine to the party.

(2) Mary did not bring both beer and wine to the party.

The traditional account maintains that the content of scalar inferences are derived through a combination of general reasoning, certain norms of conversation, and lexical scales (Horn, 1972; Grice, 1975). These lexical scales are proposed to be made up of multiple items that are ordered by informational strength (e.g., <some, many, most, all>, <or, and>, <might, must>). When a sentence includes an item from this scale, as in the disjunctive statement in (1), hearers often derive scalar inferences by negating the informational content of certain alternative sentences. These alternative sentences are formed by replacing the item that was

[1] Also known as *scalar implicatures*.

used (e.g., disjunction) with one of its stronger scale-mates (e.g., conjunction). On the traditional account, when a hearer is presented with a sentence like (1), they reason that if the speaker could have produced the stronger assertion in (3), then they would have. The fact the speaker did not produce the sentence in (3) leads the hearer to conclude that the speaker believes it to be false. In order for the hearer to align her discourse model as closely as possible to that of the speaker, the hearer enters the negation of the stronger statement into her discourse model. This explains the derivation of the scalar inference in (2).

(3) Mary brought beer and wine to the party.

Contemporary accounts of scalar inference derivation vary in how closely they resemble the traditional account. One contemporary account, known as the *Pragmatic account*, is a straightforward extension of the traditional account. Like the traditional account, the Pragmatic account views the derivation of scalar inferences as driven by domain-general reasoning and conversational norms(Spector, 2003; Sauerland, 2004; Geurts, 2010). As a result, the Pragmatic account supposes that scalar inferences are derived during the pragmatic stage of sentence interpretation, after the semantic content of a sentence has been computed. Another contemporary account of scalar inferences is known as the *Grammatical account*. The Grammatical account proposes that scalar inferences are derived within language's recursive computational system, rather than at a later, pragmatic stage of sentence interpretation. According to the Grammatical account, scalar inferences are derived through the application of a silent grammatical operator (sometimes labelled as 'EXH', which is shorthand for 'exhaustify') (Chierchia, 2004, 2006; Fox, 2007; Chierchia et al., 2011; Chierchia, 2013). For our purposes, it is sufficient to define 'EXH' as a command to apply the traditional scalar inference process to the linguistic content in its scope.[2] Take sentence (4), for example. Two inferences have been associated with a sentence like (4); namely, the inferences indicated in (5) and (6).

(4) Every woman brought beer or wine to the party.

(5) Not every woman brought both beer and wine to the party.

[2]For a more comprehensive and formal description, see Chierchia (2006) and Fox (2007).

(6) None of the women brought both beer and wine to the party.

The Grammatical account takes the inferences in both (5) and (6) to be derived through the same scalar inference process (i.e. EXH), applied to different parts of the assertion in (4). By contrast, while the Pragmatic account can easily capture the inference in (5) as a scalar inference, the same cannot be said for the inference in (6). On that account, the inference in (6) is captured, either through a more complex instantiation of the scalar inference process (e.g., Spector 2003; Sauerland 2004; Chemla 2009a), or as a *marginal* interpretation computed through an alternative mechanism altogether (e.g., Geurts and van Tiel (2013)).

One way to adjudicate between these accounts is to investigate language users' behaviour with inferences on which there is disagreement. For example, one could attempt to see whether there is a continuity in language users' behaviour between the inference in (5), and the inference in (6). Such behaviour would be evidence in favour of the Grammatical account, given that it takes both inferences to be derived through the same process. By contrast, while such an outcome does not run counter to the the Pragmatic account's predictions, unlike the Grammatical account, it would also be consistent with other outcomes, making it less directly falsifiable.

In addition to the considerable theoretical work on scalar inferences, a great deal of experiment-based research has been conducted on the scalar inferences of adults and children. The present thesis will mainly focus on children's behaviour with scalar inferences. The initial research findings in this area suggested that children derived scalar inferences at quite a low rate, as compared to adults (Noveck 2001; Chierchia et al. 2001; Gualmini et al. 2001; Papafragou and Musolino 2003; Guasti et al. 2005; Barner et al. 2011; Foppolo et al. 2012, among others). However, much of this early work was focused on quite a small number of 'classical' scalar inferences. More recent work has investigated children's abilities with a wider range of inferences. This research had uncovered a number of inferences that are amenable to a scalar inference analysis, which children derive readily (Stiller et al., 2015; Tieu et al., 2016; Hochstein et al., 2016; Singh et al., 2016). This variable success in scalar inference computation presents a challenge for theories attempting to explain children's behaviour in this area.

One quite successful attempt to account for this variation is presented by a group of explanations which we call the *Alternatives-based approach*. The Alternatives-based approach captures children's behaviour with scalar inferences by attributing to children limitations that affect their ability to generate the alternative sentences from which scalar inferences are derived (Chierchia et al., 2001; Gualmini et al., 2001; Reinhart, 2006; Barner and Bachrach, 2010; Barner et al., 2011; Tieu et al., 2016; Singh et al., 2016). Crucially, for reasons that we will detail later, the scalar inferences that children have proven able to derive are plausibly not affected by these limitations. In this way, the Alternatives-based approach is able to account for both children's documented difficulties with scalar inference derivation, as well as their successes. In addition to accounting for these previous findings the Alternatives-based approach makes predictions regarding children's behaviour with scalar inferences that have not yet been investigated. Specifically, according to the Alternatives-based, if the linguistic material from which a scalar inference's alternative sentences are composed is contained in the asserted sentence, then children will successfully derive that inference.

This thesis presents three studies that investigate the behaviour of adults and children towards a series of scalar inferences that have not been investigated previously, and which the Alternatives-based approach predicts children will be successful in deriving. The results of these investigations contribute to our understanding of both the nature of scalar inferences, as well as how children acquire and develop an understanding of them. Specifically, the results of these studies are largely consistent with the predictions of the Alternatives-based approach, and so support it as a viable explanation of children's variable success with scalar inference computation. Moreover, our results help to adjudicate between the Pragmatic and Grammatical accounts of scalar inferences, as well as providing insights into several other important theoretical issues.

1.1.1 Different theories of scalar inference derivation

As noted earlier, the Pragmatic account and the Grammatical account of scalar inferences appeal to different mechanisms. The Pragmatic account takes the derivation of scalar inferences to be a result of general reasoning and conversational norms, whereas the Grammatical account takes it to be a result of a silent grammatical operator. These mechanisms differ in how they apply to sentences. The mechanisms appealed to by the Pragmatic account are applied at the pragmatic stage of language processing, after the semantic content of a sentence has been computed. As a result, the mechanisms invoked by the Pragmatic account can only apply to whole sentences. In contrast, the grammatical operator appealed to by the Grammatical account is part of the recursive computational system of sentence interpretation. Therefore, this operator can be applied to different parts of a sentence, as part of this computational process. The differences between these accounts means that they have different perspectives on which inferences should be counted as scalar inferences. For example, consider the sentence in (7). According to the Grammatical account, the EXH operator can be applied at two sites; at the whole sentence level, as in (8), and embedded under the universal quantifier, as in (9).

(7) Every woman brought beer or wine to the party.

(8) EXH[Every woman brought beer or wine to the party.]

(9) Every woman EXH[x brought beer or wine to the party.]

Application of EXH at these different sites results in the derivation of different inferences. The inference in (10) results when EXH applies to the whole sentence, as indicated in (8). The inference in (11) results when EXH is embedded under the universal quantifier, as indicated in (9). Again, according to the Grammatical account, both inferences are derived using the same scalar inference process. The difference is just that the operator EXH is being applied to different parts of the asserted sentence.

(10) Not every woman brought both beer and wine to the party.

(11) None of the women brought both beer and wine to the party.

In contrast, the Pragmatic account takes (10) to be a scalar inference, but it is not able to explain the inference in (11) using exactly the same process. Instead, there are two ways the Pragmatic account captures inferences like (11). One proposal introduces greater complexity into the scalar inference process so that the sentence in (12) can be included as an alternative sentence of the assertion in (7) (Spector, 2003; Sauerland, 2004; Chemla, 2009a; Chemla and Spector, 2011).

(12) Some woman brought beer and wine to the party.

Specifically, in order to include such an alternative two assumptions need to be added to the scalar inference process. First, it must be assumed that alternative sentence generation is able to involve the replacement of multiple lexical items. This is because the sentence in (12) is generated from the asserted sentence in (7) through the replacement of both *every* with *some* as well as *or* with *and*. Second, it must be assumed that alternative sentences can be logically independent from asserted sentences, rather than needing to be stronger than (i.e. asymmetrically entailing) asserted sentences. This is because the alternative sentence in (12) is logically independent from the asserted sentence in (7). Accepting these two assumptions allows these versions of the Pragmatic account to capture both the inference in (10) and the inference in (11) through the same scalar inference process. Even with these assumptions, however, the derivation of the inference in (11) is more complex than the inference in (10), because it requires the replacement of multiple lexical items and a more nuanced relationship between the alternative sentence and the asserted sentence.

The second way that the Pragmatic account has attempted to capture the inference in (11) is by appealing to an alternative mechanism. One example of this strategy is presented in Geurts and van Tiel (2013), which attempts to explain the derivation of this inference by appealing to pragmatics-based *truth-conditional narrowing*. According to this proposal, a hearer is sometimes led by certain contextual cues (e.g., prosodic stress) to take a speaker to be attaching a non-conventional meaning to a word. For example, a hearer could take the sentence in (13), where the second *drinks* is given contrastive stress, to be communicating something beyond its conventional meaning - such as, *Mary drinks a lot of alcohol*.

(13) When Mary drinks, she DRINKS.

Geurts et al. contend that the same process can be applied to the sentence in (7), through the stressing of the disjunctive operator *or*, to derive the inference in (11). Therefore, whether it is through a more complex instantiation of the scalar inference process, or a different mechanism altogether, the Pragmatic account takes there to be some difference in the processes by which the inference in (10) and the inference in (11) are derived. The upshot is that while both the Pragmatic account and the Grammatical account explain how language users derive the inferences in (10) and (11) from an assertion like (7), there is a greater continuity in the Grammatical account's derivation of the two inferences.[3]

Another case where these accounts differ is in their classification of the inferences associated with sentences like (14) - namely, inferences like (15). One way to account for such an inference is through the recursive application of the scalar inference process, embedded under the universal quantifier, as in (16).

(14) Every woman is allowed to bring beer or wine to the party.

(15) Every woman is such that she is allowed to bring beer to the party and she is allowed to bring wine to the party.

(16) Every woman EXH[EXH[x is allowed to bring beer or wine to the party.]]

While both the Pragmatic account and the Grammatical account can apply the scalar inference process recursively (Spector, 2003; Fox, 2007), it is only according to the Grammatical account that the scalar inference process can be embedded. Therefore, only the Grammatical account classifies the inference in (15) as a scalar inference.[4] In summary, the Pragmatic account and Grammatical account adopt different categorisations for a number of inferences. Only the Grammatical account takes the inferences in (10), (11), and (15) to be derived through exactly the same scalar inference process. By extension, the Grammatical account predicts a greater continuity in language users' behaviour towards these different inferences.

In the present thesis, we test the predictions made by these competing scalar inference accounts by investigating adults' and children's behaviour with the inferences we have just described. Chapter 4 investigates behaviour by adults and

[3]An anonymous examiner noted that, as with the Pragmatic account, the Grammatical account could also be construed as deriving the inferences in (10) and (11) via somewhat different processes. That is, according to the Grammatical account they are each derived through the application of the EXH operator at difference scope sites. While we agree with this observation, we would still assert that there is an even greater difference between the mechanisms appealed to by the Pragmatic account to account for the two inferences in (10) and (11). This is especially the case for the version of the Pragmatic account presented by Geurts and van Tiel (2013), which appeals to different mechanisms altogether to capture each of these inferences.

[4]See Schulz (2006) and Geurts (2010) for explanations of the inference in (15) that appeal to alternative pragmatic mechanisms.

by children in computing the inferences in (10) and (11). Chapter 3 investigates the inference in (15). To forestall suspense, the findings of these investigations reveal that language users' behaviour are largely consistent with the Grammatical account; these inferences would appear to be derived through the same process.

1.1.2 Children's behaviour with scalar inferences

Children's behaviour with scalar inferences has been the focus of a series of experimental investigations for nearly two decades (e.g., Chierchia et al. 2001; Noveck 2001; Papafragou and Musolino 2003; Guasti et al. 2005; Foppolo et al. 2012; Bill et al. 2016, among many others). The most consistent result to come out of the earlier work in this area was that children struggled to derive scalar inferences, particularly compared to adults. A typical example of a study producing this kind of result is Noveck (2001). Noveck investigated children's interpretations of sentences like (17), which have been associated with inferences like (18).

(17) Some giraffes have long necks.

(18) Not all giraffes have long necks.

Noveck carried out this investigation by employing a *reasoning scenario*, wherein participants were presented with test sentences like (17) and asked whether they agreed with them or not. Crucially, if participants derived the relevant inference in (18), then their interpretation would not be consistent with their world knowledge that giraffes have long necks. Therefore, a participant rejecting a test sentence was taken as evidence of that participant deriving the scalar inference. Noveck included 31 8-year-olds, 30 10-year-olds, and 15 adults in the experiment. Noveck's results suggested that adults were deriving inferences like (18) 69% of the time, while the two child groups derived these inferences less often (between 11% of the time (8 y/o), and 15% of the time (10 y/o)). This result is representative of the bulk of the earlier findings from research investigating children's abilities with scalar inferences.

Despite the bulk of early results following this pattern, these investigations were focused on inferences associated with a relatively small number of 'classical' scales (e.g., 'or/and', 'some/all', 'might/must'). As a result of the theoretical literature proposing a scalar inference analysis for a wider range of inferences, more recent work expanded the number and kinds of inferences that have been investigated. This work has found that there are a handful of inferences with a scalar inference analysis that children are able to successfully derive (Papafragou and Musolino, 2003; Barner and Bachrach, 2010; Barner et al., 2011; Huang et al., 2013; Stiller et al., 2015; Tieu et al., 2016; Hochstein et al., 2016; Singh et al., 2016). One study which produced such a result was done by Tieu et al. (2016), and was investigating children's interpretations of sentences like (19). Such sentences have been associated with *free choice inferences* like (20), which have an established scalar inference analysis (Kratzer and Shimoyama, 2002; Alonso-Ovalle, 2005; Fox, 2007; Chierchia, 2013).

(19) Kung Fu Panda may push the green car or the orange car.

(20) Kung Fu Panda may push the green car and Kung Fu Panda may push the orange car.

Tieu et al. (2016) used the *Truth Value Judgment Task* (Crain and Thornton, 1998) to investigate children's derivation of free choice inferences. This task involved participants being presented with stories, and then being presented with a description of those stories from a puppet. The puppet's description included a test sentence like the sentence in (19). The participant is then asked to judge whether the puppet's description was right or wrong. In Tieu et al. the stories were inconsistent with the relevant free choice inference (e.g., Kung Fu Panda was only allowed to push the green car). Children who computed such inferences, therefore, were expected to reject the test sentences. Tieu et al. tested 22 Mandarin-speaking children using Mandarin translations of target sentences like (19). Tieu et al. found that children derived free choice inferences 91% of the time, compared to a rate of 18% for a more traditional scalar inference (i.e. the exclusivity inference).

Tieu et al. (2016)'s finding, of children deriving free choice inferences readily has been replicated for a number of other scalar inferences, including *exactly-n inferences* (Papafragou and Musolino, 2003; Barner and Bachrach, 2010; Huang et al., 2013), *ignorance inferences* (Hochstein et al., 2016), *conjunctive inferences* (Singh et al., 2016; Tieu et al., 2017), and *ad-hoc inferences* (Stiller et al., 2015).

In sum, the bulk of early work found that children struggled to derive scalar inferences (e.g., Noveck 2001; Papafragou and Musolino 2003; Guasti et al. 2005). However, a handful of more recent investigations have found that there are certain scalar inferences (e.g., free choice inferences) that children derive successfully. This pattern of results is a challenge for theories attempting to capture children's abilities with scalar inferences. We will now consider some of the different attempts that have been made in this regard.

1.1.3 Explaining children's behaviour with scalar inferences

There have been a variety of explanations proposed which attempt to account for children's behaviour with scalar inferences. These explanations have largely focused on explaining the more common behaviour of children deriving scalar inferences at a low rate. However, some also offer insights into why children's abilities with scalar inferences vary between different scalar inferences. We will group these explanations into two distinct approaches, based on the nature of the limitation they attribute to children.

The Pragmatics-based approach

First, we will consider a group of explanations that we will refer to as the *Pragmatics-based approach*. This family of theories attribute children's behaviour with scalar inferences to limitations in their pragmatic knowledge (Noveck, 2001; Katsos and Bishop, 2011; Skordos and Papafragou, 2016). For example, Katsos and Bishop (2011) propose that children are more *pragmatically tolerant* than adults. This means that children are less likely to reject a sentence based on 'pragmatic' content (such as scalar inferences) being false in the discourse context. In support of this proposal Katsos et al. show that when children are given an intermediate judgment option, they will choose it, rather than accepting or rejecting a test sentence whose scalar inference is inconsistent with the discourse context. Katsos et al. conclude that children's documented low rate of scalar inference derivation is at least partly caused by the relevant experiments not included this kind of intermediate judgment option, resulting in children's abilities being perceived as worse than they really are. In this way Katsos et al. are able to account for children's recorded difficulties with scalar inference derivation. However, Katsos et al. do not provide a clear account of why more recent results have found that children derive certain scalar inferences (e.g., free choice inferences) readily, even though they also employed a binary judgment task. That is, presumably this pragmatic tolerance should consistently affect children's derivation of scalar inferences, and so, without further amendments, Katsos et al. do not capture these successes.

Another explanation within the Pragmatics-based approach was recently presented in a paper by Skordos and Papafragou (2016). Skordos et al. propose that at least some of children's variable behaviour with scalar inferences can be attributed to them having a limited understanding of how the context affects which alternative sentences are considered 'relevant'.[5] In support of this proposal, Skordos et al. present a series of experiments in which they manipulate elements of the context to increase or decrease the saliency of various alternative sentences. The results of these experiments suggest that the rate at which children derive scalar inferences changes, based on how salient the alternative sentences are perceived in a given context. In this way, Skordos et al. go further than Katsos and Bishop (2011) because, in addition to accounting for children's low rates of scalar inference derivation, the mechanism they propose (i.e., relevance of alternative sentences) expects more variability in children's behaviour. Having said that, it would appear that such a mechanism can only capture some of children's variable behaviour. That is, studies like Tieu et al. (2016) and Hochstein et al. (2016) found that children derived certain scalar inferences (e.g., free choice inferences) at a much higher rate than more 'classical' scalar inferences (e.g., exclusivity inference), even though the inferences were presented in equivalent contexts. Given that Skordos et al.'s proposal is based on contextual differences, it is not clear that it can account for this documented variability. Therefore, the explanation proposed by Skordos et al. appears only able to capture some of the variation with scalar inferences exhibited by children.[6]

The Pragmatics-based approach could account for children's variable performance by suggesting that the inferences children derive readily (e.g., free choice in-

[5] Note that a similar, although less developed, idea was proposed by Noveck (2001).

[6] In Skordos and Papafragou (2016)'s defence, they never actually claim that their proposal can account for all of children's behaviour with scalar inferences. In fact, they suggest that their proposed limitation may be 'just' one of several limitations affecting children's behaviour with scalar inferences.

ferences) are not derived as scalar inferences at all. That is, the Pragmatics-based approach could adopt or create an alternative non-scalar inference analysis for these inferences, thereby removing them from the group of inferences their approach is designed to account for. However, such a move is not 'cost-free'. That is, in addition to being patently post-hoc, adopting such a position would mean losing the gains in parsimony achieved by explaining how so many, seemingly disparate inferences, are derived, through the one scalar inference process. Given this cost, an explanation that could account for children's behaviour while retaining a scalar inference analysis for the relevant inferences is preferable. With this is in mind, we turn to the second group of explanations attempting to account for children's behaviour with scalar inferences.

The Alternatives-based Approach

We refer to the second group of explanations as the *Alternatives-based approach*. The Alternatives-based approach attributes children's variable behaviour with scalar inferences to limitations affecting children's ability to generate the alternative sentences involved in scalar inference derivation (Chierchia et al., 2001; Gualmini et al., 2001; Reinhart, 2006; Barner and Bachrach, 2010; Barner et al., 2011; Tieu et al., 2016; Singh et al., 2016). The explanations that make up the Alternatives-based approach can be broken down into at least two distinct proposals.

One proposal is that children's behaviour with scalar inferences is a result of limitations in their processing abilities (e.g., working memory capacity) (Chierchia et al., 2001; Gualmini et al., 2001; Reinhart, 2006; Tieu et al., 2016). The other proposal is that children are limited in their knowledge of the lexical scales involved in scalar inference derivation (Barner and Bachrach, 2010; Barner et al., 2011). These proposals are similar in that both proposed limitations are expected to negatively affect children's ability to generate the same alternative sentences, and their associated scalar inferences. Specifically, these limitations are expected to affect children's generation of alternative sentences composed through lexical replacement. One such alternative sentence is the sentence in (21). The sentence in (21) is generated from the asserted sentence in (22) and is involved in the derivation of the exclusivity inference in (23). The sentence in (21) is generated from the sentence in (22) through lexical replacement - namely, by replacing the disjunctive lexical item ‘or’ with its stronger scale-mate, the conjunctive lexical item ‘and’.

(21) Mary brought beer and wine to the party.

(22) Mary brought beer or wine to the party.

(23) Mary did not bring both beer and wine to the party.

Generating alternative sentences through lexical replacement is plausibly quite taxing on processing resources and clearly requires knowledge of the relevant lexical scale to accomplish. Given that, according to the Alternatives-based approach, one or both of these are limited in children, this approach expects for children to struggle to derive scalar inferences that are associated with alternatives composed through lexical replacement, such as the exclusivity inference. As most of the ‘classical’ scalar inferences are derived through this process, the Alternatives-based approach is able to account for the early results suggesting that children have difficulties deriving such inferences.

The Alternatives-based approach also provides an explanation of why, in certain cases, children have been found to derive scalar inferences successfully. Namely, the presentation of certain material in in the environmental or linguistic contexts, plausibly reduces the processing and scale knowledge requirements associated with scalar inference derivation. For example, from asserted sentences like (24), free choice inferences like (25) are proposed to be derived from the alternative sentences in (26) and (27) (Fox, 2007). Notably, both these alternative sentences can be formulated from linguistic material contained in the asserted sentence in (24). Presenting this linguistic material in the assertion means that language users do not need to retrieve it from a their mental lexicons, thereby reducing the processing and scale knowledge requirements of deriving the associated inference.

(24) Mary is allowed to bring beer or wine to the party.

(25) Mary is allowed to bring beer to the party and Mary is allowed to bring wine to the party.

(26) Mary is allowed to bring beer to the party.

(27) Mary is allowed to bring wine to the party.

Considering the scalar inferences children have been found to derive readily, many of them are associated with asserted sentences which have this property (e.g., free choice inferences (Tieu et al., 2016), ignorance inferences (Hochstein et al., 2016), conjunctive inferences (Singh et al., 2016; Tieu et al., 2017)), and ad-hoc inferences[7]. In this way, the Alternatives-based approach is able also to account for children's successes in scalar inference derivation.

In summary, like the Pragmatics-based approach, the Alternatives-based approach can account for results showing that children struggle to derive scalar inferences. In contrast to the Pragmatics-based approach, however, the Alternatives-based approach also accounts for many of the cases where children have been found to be successful in computing scalar inferences.

We should note that there are scalar inferences that children have been found to derive readily, which even the Alternatives-based approach struggles to account for. For example, when presented with sentences containing numerals, like (28), children have been found to readily derive exactly-n inferences like (29) (Papafragou and Musolino, 2003; Barner and Bachrach, 2010; Huang et al., 2013). Unlike free choice inferences, the linguistic material from which the relevant alternative sentence in (30) is composed, is not a part of the asserted sentence in (28).

(28) Mary brought one beer.

(29) Mary brought exactly one beer.

(30) Mary brought two beers.

Therefore, children's successful derivation of exactly-n inferences cannot be accounted for in the same way as their derivation of free choice inferences. As a result, even the Alternatives-based approach may need to adopt an alternative non-scalar inference analysis for such inferences, in order to account for children's behaviour.[8]

[7]The study by Stiller et al. (2015), which found children successfully deriving ad-hoc inferences presented graphical representations of the relevant alternatives and their scalar relationship in the environmental context. Presenting such material conceivably produces a similar effect in enabling children to overcome the Alternative-based approach's proposed limitations. See Chierchia et al. (2001); Gualmini et al. (2001) and Barner et al. (2011) for similar cases.

[8]Having said that, it could be that the exactly-n inference is an exceptional case and could

To summarise, in contrast to the Pragmatics-based approach, the Alternatives-based approach is able to account for both the typically low rate at which children have been found to derive most traditional scalar inferences, while also explaining why, in certain cases, children have been found to succeed. Moreover, the Alternatives-based approach makes clear predictions regarding how children will perform with many scalar inferences that have not yet been investigated. That is, if the linguistic material from which a scalar inference's alternatives are composed is contained in the asserted sentence, then children are expected to successfully derive that inference. One way then, to test the Alternatives-based approach is to investigate children's behaviour with such scalar inferences to see whether it is consistent with these predictions. The current thesis presents a series of such investigations.

1.2 Chapter summary

The current thesis presents a series of chapters investigating the theoretical issues we have discussed thus far. Note that as these chapters are based on stand-alone papers there is some repetition between them. We will now give a brief outline of each of these chapters.

1.2.1 Chapter 2

Chapter 2 investigates children's interpretations of sentences like (31), which have been associated with *distributive inferences* like (32). The distributive inference is indicated in (32). This inference has been analysed as a scalar inference in Crnič et al. (2015). Another inference that has been associated with sentences like (31) is the *conjunctive inference* in (33) (Singh et al., 2016).

(31) Every woman brought beer or wine to the party.

be accounted for along the following lines (originally proposed by Barner and Bachrach (2010)). Children are explicitly presented with numbers as a part of an ordered lexical scale (i.e., <1, 2, 3, etc.>) from a very young age (Fuson, 1988). As a result children's understanding of this lexical scale is more developed than most others, making the scalar inferences derived from it (i.e., the exactly-n inference) immune from the limitations proposed by the Alternatives-based approach. Note that such an explanation could not necessarily be extended to other cases in which children's behaviour went against the Alternatives-based approach's expectations.

(32) At least one woman brought beer to the party and at least one woman brought wine to the party.

(33) Every woman is such that she brought beer and wine to the party.

Similar to free choice inferences, the linguistic material from which both these inferences' alternatives are composed is contained in the asserted sentence. As a result, the Alternatives-based approach expects children to be successful in deriving one of these inferences from sentences like (31). We test this prediction by comparing the rates at which adults and children derive an inference-based interpretation (based on deriving either a distributive inference or a conjunctive inference). The findings of this experiment are in line with the Alternatives-based approach's prediction, with children accessing an inference-based interpretation at an adult-like rate.

1.2.2 Chapter 3

Chapter 3 investigates children's behaviour with two kinds of free choice inference. The sentence in (34) is associated with the *basic free choice inference* in (35), and the sentence in (36) is associated with the *universal free choice inference* in (37) (Kratzer and Shimoyama, 2002; Alonso-Ovalle, 2005; Fox, 2007; Chierchia, 2013). For both these types of free choice inference, the linguistic material from which their alternative sentences are composed is contained in the relevant asserted sentence. Therefore, the Alternatives-based approach expects children to be successful in deriving both the basic free choice inference and the universal free choice inference.

(34) Mary is allowed to bring beer or wine to the party.

(35) Mary is allowed to bring beer to the party and Mary is allowed to bring wine to the party.

(36) Every woman is allowed to bring beer or wine to the party.

(37) Every woman is allowed to bring beer to the party and every woman is allowed to bring wine to the party.

The Grammatical account of scalar inferences classifies both these kinds of free choice inferences as scalar inferences. Therefore, the Grammatical account expects that children's abilities with these inference will mirror that of other equivalent scalar inferences. That is, assuming the Alternatives-based approach is correct, the Grammatical account predicts that children will be successful in deriving both of these free choice inferences.

We test this shared prediction of the Alternatives-based approach and the Grammatical account by investigating adults' and children's performance with basic free choice inferences and universal free choice inferences. The results from this experiment are in the direction predicted by these theories, with children deriving both free choice inferences at a high rate, relative to previous results with more traditional scalar inferences. Despite children's high rate of derivation, adult participants were found to derive these free choice inferences at a higher rate still, a difference that the Alternatives-based approach does not account for, and which requires further explanation.

1.2.3 Chapter 4

Chapter 4 explores children's behaviour with two more inferences that are both associated with *EverySome sentences* like (38). One of these inferences we call the *NotEvery inference*. This inference is indicated in (39). In addition, sentence (38) is associated with what we call the *None inference*, as indicated in (40).

(38) Every woman brought some of her beers.

(39) Not every woman brought all of her beers.

(40) None of the women brought all of her beers.

As with the scalar inferences investigated in the previous chapters, the linguistic material from which these inferences' alternative sentences are composed is contained in the asserted sentence. As with the other scalar inferences we investigated then, the Alternatives-based approach predicts that children will be successful in deriving these inferences. Specifically, the Alternatives-based approach predicts that children will be successful in deriving an inference-based interpretation of asserted sentences like (38) (i.e. that they will derive either the NotEvery inference or the None inference).

Chapter 4 also tests a theory presented in Crain et al. (1994) related to the considerations that are proposed to guide adults and children when they are interpreting sentences. In brief, Crain et al. propose that children are guided by learnability considerations to prefer stronger interpretations, while adults are guided by parsing considerations to prefer weaker interpretations. This proposal predicts that when accessing an inference-based interpretation of EverySome sentences like (38), children will prefer interpretations including the stronger None inference in (40), while adults will prefer interpretations including the weaker NotEvery inference in (39).

Finally, Chapter 4 also contributes to the debate between competing accounts of scalar inferences. While the NotEvery inference is straightforwardly captured as a scalar inference by both the Grammatical and Pragmatic accounts, only the Grammatical account captures the None inference through exactly the same process. Therefore, our investigation in Chapter 4 provides another test of the Grammatical accounts expectation of continuity in language users' behaviour with inferences it classifies as scalar inferences. Including the expectations of the Alternatives-based approach and of Crain et al. (1994), this equates to the Grammatical account expecting children to prefer interpretations based on having derived the None inference over the NotEvery inference.

In Chapter 4, we test the predictions of these different theories by investigating adults' and children's interpretations of EverySome sentences like (38). The results of this investigation are consistent with each of the mentioned theories. Children access an inference-based interpretation of the target sentences at an adult-like rate, as expected by the Alternatives-based approach. Moreover, the inference on which these interpretations were based differed between groups in the manner predicted by Crain et al. (1994). Finally, children were found to prefer the None inference over the NotEvery inference, as expected by the Grammatical account.

1.3 Conclusion

The present thesis reports the findings of three experimental investigations of children's abilities with different inferences. These investigations are primarily focused on testing the predictions of the Alternatives-based approach, however, they also provided the opportunity to test and contribute to a number of other theoretical discussions in this area. The findings of these investigations are consistent with the Alternatives-based approach, providing further support for it as a viable account of children's variable success in computing scalar inferences. Moreover, the findings are in line with the expectations of the Grammatical account of scalar inference derivation, supporting it as a description of the scalar inference process. Finally, the results are consistent with the expectations of Crain et al. (1994)'s proposal regarding interpretation strategies, and so provide further support for this proposal as an accurate account of age-based differences in sentence interpretation.

The findings of the present thesis contribute to our understanding of the nature of scalar inferences, how children acquire and develop an understanding of these inferences, as well as how they are interact with general principles of sentence interpretation and language learnability.

Chapter 2

Children's distributive inferences

This chapter is based on the following paper:

Pagliarini, E., **Bill, C**., Romoli, J., Tieu, L., and Crain, S. (2017). On children's variable success with scalar inferences: Insights from disjunction in the scope of a universal quantifier. Under review.

Author contributions:[1]
All authors assisted in the conception and design of the experiment. EP and **CB** created the materials for the experiment, with feedback from SC. **CB** conducted the data collection, with assistance from EP. EP conducted the statistical analysis, assisted by **CB** and LT. EP drafted the results section of the paper, all other sections of the paper were drafted by **CB**. Paper revisions were contributed to by all authors by providing feedback, as well as by directly editing the paper.

[1]Authors referred to by their initials

Abstract

Previous developmental studies have revealed variation in children's ability to compute scalar inferences. While children have been shown to struggle with standard scalar inferences (e.g., from scalar quantifiers like *some*) (Noveck, 2001; Chierchia et al., 2001; Papafragou and Musolino, 2003; Guasti et al., 2005), there is also a growing handful of inferences that children have been reported to derive readily (Papafragou and Musolino, 2003; Barner and Bachrach, 2010; Stiller et al., 2015; Tieu et al., 2016; Hochstein et al., 2016). One recent approach, which we refer to as the *Alternatives-based approach*, attributes the variability in children's performance to limitations in how children engage with the alternative sentences that are required to compute the relevant inferences. According to this approach, if the alternative sentences are a component part of the test sentences, children should be better able to compute the inference. In this paper, we investigated this prediction by assessing how children and adults interpret sentences that embed disjunction under a universal quantifier, such as *every elephant caught a big butterfly or a small butterfly*. These sentences give rise to the *distributive inference* that *some elephant caught a big butterfly and some elephant caught a small butterfly* (Gazdar, 1979; Fox, 2007; Crnič et al., 2015). It has been argued that this inference is derived using alternatives that are contained within the test sentence, making these sentences an ideal test case for evaluating the predictions of the Alternatives-based approach. The findings of our experimental study reveal that children are indeed able to successfully compute distributive inferences at adult-like rates; moreover, we also observe evidence for the presence of a *conjunctive inference*, i.e. *every elephant caught both a big butterfly and a small butterfly*, previously reported in studies by Singh et al. (2016) and Tieu et al. (2017). Both of these findings are in line with the predictions of the Alternatives-based approach, and provide support for it as a viable explanation of children's variable success in computing scalar inferences.

2.1 Introduction

According to the standard semantics of disjunction, which is derived from classical propositional logic, a basic disjunctive sentence like (1) is true if at least one of its disjuncts is true. However, when a sentence like (1) is used in everyday conversation, it often gives rise to additional inferences. In addition to the literal inclusive meaning, language users often also derive the inferences in (2) and (3) (Gazdar 1979; Sauerland 2004; Fox 2007, among many others). These inferences are known as *scalar inferences*.[2]

(1) The elephant caught a big butterfly or a small butterfly.

(2) The elephant didn't catch both a big butterfly and a small butterfly.

(3) The speaker is ignorant as to whether the elephant caught a big butterfly and as to whether the elephant caught a small butterfly.

The traditional Gricean explanation for how such inferences are derived involves a combination of general reasoning and assumptions about how rational agents interact in conversation (Grice 1975; Gamut 1991, among many others). More specifically, the proposal is that the hearer assumes that at any given time during a conversation, the speaker will produce the most informative utterance that is relevant for the purposes of the conversation and that she believes to be true. For example, when the speaker utters the sentence in (1), the hearer will reason that if the speaker could have uttered the more informative sentence in (4), then she would have done so.

(4) The elephant caught a big butterfly and a small butterfly.

The fact that the speaker uttered the sentence in (1), rather than the sentence in (4), leads the hearer to infer that the speaker does not have sufficient evidence that the sentence in (4) is true. If the hearer takes the speaker to be well-informed, then the hearer is led to infer the negation of the sentence in (4), which is the *exclusivity inference* in (2).[3]

[2] Also known as *scalar implicatures*.

[3] See Geurts (2010) for an alternative, non-scalar inference analysis of the exclusivity inference.

The question of how such inferences are derived has been subject to much subsequent work since Grice and is the subject of ongoing debate. Approaches differ in whether the source of the phenomenon lies within the grammar or in a post-grammatical pragmatic domain. One aspect that most approaches agree on is that the hearer's reasoning about what the speaker might have said should be constrained to a set of alternatives (Horn, 1972; Gazdar, 1979; Katzir, 2007; Fox and Katzir, 2011). A standard idea in the literature is that the alternatives to what the speaker said are generated on the basis of certain words being associated with others in the language user's mental lexicon. For example, Horn (1972) suggests that certain words occupy positions on lexical scales. The scales are ordered by informational strength (e.g., <or...and>, <some...all>, <might...must>).[4] When a speaker utters a sentence that includes one of the terms on the scale, alternative sentences are generated by replacing the relevant lexical item with one of its stronger scale-mates. For instance, a stronger alternative to an utterance like (1) is the sentence in (4). The alternative sentence in (4) is generated by replacing the lexical item 'or' in (1) with its more informative scale-mate 'and'. Because the hearer supposes that the speaker is being cooperative and has produced the strongest statement that she was prepared to make, the hearer infers that the sentence in (1), with 'or', is the strongest statement the speaker felt she had evidence for. Therefore, the hearer infers the negation of the stronger alternative sentence with 'and'. The negation of the alternative sentence in (4) yields the exclusivity inference in (2).[5]

The relationship between the exclusivity inference and the *ignorance inference* in (3) is controversial (for discussion, see Sauerland 2004; Fox 2007; Meyer 2013). One approach invokes the same mechanism that underlies the derivation of scalar inferences. On this approach, the ignorance inference in (3) involves a comparison of the speaker's utterance in (1) with the alternative sentences in (5) and (6).

[4]The informativity of a given sentence is based on entailment relations between it and other sentences. Sentence A is more informative than sentence B iff A asymmetrically entails B. For example, the sentence in (4) is more informative than the sentence in (1) because (4) entails (1), but (1) does not entail (4).

[5]More recent work on alternatives criticises this approach for merely stipulating which terms are associated with others in the lexicon, and argues instead for more general alternative-generating algorithms for any given sentence. For discussion, see Katzir (2007); Fox and Katzir (2011); Trinh and Haida (2015), and Breheny et al. (2016).

(5) The elephant caught a big butterfly.

(6) The elephant caught a small butterfly.

These alternatives correspond to the individual disjuncts of the sentence in (1) (Gazdar, 1979; Sauerland, 2004). The alternative sentences in (5) and (6) are both more informative than the asserted sentence in (1). If the speaker had sufficient evidence that the sentences in (5) and (6) were true, then she should have uttered these alternatives instead of the assertion in (1). The fact that the speaker chose not to say the alternative sentences in (5) or (6) leads the hearer to infer that the speaker does not have sufficient evidence that the sentences in (5) or (6) are true, generating the ignorance inference in (3).[6]

Note that there is no lexical replacement involved in generating the ignorance inference, because the alternatives in (5) and (6) correspond to the individual disjuncts of the asserted sentence in (1) (Gazdar, 1979; Sauerland, 2004).

The exclusivity and ignorance inferences can both be derived through reasoning over alternative sentences that the speaker might have uttered instead of what she actually said. Because of this, both inferences are considered to be scalar inferences on some accounts. However, for reasons that will become clear later, what is most relevant for our purposes is a distinction in how the alternatives for each inference are generated; one involves lexical replacement (i.e. the exclusivity inference), whereas the other involves alternatives that are contained in the original assertion (i.e. the ignorance inference).

In addition to the substantial theoretical work on inferences like these, there has also been a great deal of experimental work, including developmental studies that assess children's ability to compute such inferences. The bulk of this work has reported that children consistently derive scalar inferences at lower rates than adults do (Noveck 2001; Chierchia et al. 2001; Papafragou and Musolino 2003; Guasti et al. 2005; Foppolo et al. 2012, among many others). However, this work has primarily focused on a rather restricted set of scalar inferences, typically those derived on the basis of alternatives that are generated through lexical replacement.

More recent investigations have expanded the range of inferences investigated

[6] Unlike in the case above of the *exclusivity inference*, in this case the hearer cannot conclude that the speaker has evidence that the sentences in (5) and (6) are false, as this would lead to a contradiction with the assertion; see Sauerland (2004) and Fox (2007) for discussion.

to *plurality inferences*, *free choice inferences*, *ad hoc inferences*, *conjunctive inferences*, and ignorance inferences (Tieu et al., 2014, 2016; Stiller et al., 2015; Hochstein et al., 2016; Singh et al., 2016; Tieu et al., 2017). Some of these studies report low rates of inference derivation by children (Tieu et al. 2014), as in previous studies; however, others have reported adult-like rates of inference derivation by children (Tieu et al., 2016; Hochstein et al., 2016; Singh et al., 2016). Notably, the inferences that children have been reported to derive readily are ones that do not generate alternative sentences through lexical replacement.

One attempt to explain children's variable performance with scalar inference computation is the *Alternatives-based approach* (Reinhart, 2006; Barner et al., 2011; Tieu et al., 2016; Singh et al., 2016). On this approach, children's behaviour is explained by proposing that children suffer limitations that reduce their capacity to generate certain alternative sentences, and by extension, their related scalar inferences. Specifically, these limitations are expected to affect children's success with scalar inferences that are derived from alternative sentences composed through lexical replacement, such as exclusivity inferences. Moreover, the limitations are not expected to affect children's performance with scalar inferences that are derived from alternative sentences composed from linguistic material contained in the asserted sentence, such as ignorance inferences.

The present chapter investigates the predictions of the Alternatives-based approach by investigating children's and adults' interpretations of sentences in which a disjunction is embedded under a universal quantifier, as in (7). Such sentences constitute an ideal case study for testing the Alternatives-based approach, because the inferences that arise from such sentences require alternatives that correspond to constituents of the sentence. Notice that the sentence in (7) licenses the *distributive inference* in (8) (Gazdar, 1979; Fox, 2007; Crnič et al., 2015).

(7) Every elephant caught a big butterfly or a small butterfly.

(8) Some elephant caught a big butterfly and some elephant caught a small butterfly.

As the findings of our experimental study reveal, children appear to be adult-like in deriving such inferences. On the Alternatives-based approach, this finding is attributed to the fact that the distributive inference in (8) can be derived using alternatives that correspond to each of the disjuncts mentioned in sentence (7).

Replicating previous developmental studies of children's interpretation of disjunction (Singh et al., 2016; Tieu et al., 2017), we will also see that some children appear to compute a conjunctive inference, from sentences like (7), accessing an interpretation of them along the lines of the sentence in (9).

(9) Every elephant caught both a big butterfly and a small butterfly.

To preview, we will conclude that the results of the present investigation are in line with the predictions of the Alternatives-based approach, providing further support for it as a viable explanation of children's variable success in computing a large class of scalar inferences.

The rest of the paper is organised as follows. First, we review previous developmental studies of scalar inferences. We then outline a prediction of the Alternatives-based approach. This is followed by a description of the experiment we designed to test this prediction. We conclude with a discussion of how our results bear on the Alternatives-based approach.

2.1.1 Children's variable performance with scalar inferences

Children have been found to display variable success with scalar inference computation. We will first consider some studies that have reported non-adult-like performance, and then move on to cases in which children were adult-like.

Non-adult-like behaviour

For over a decade, developmental research has found that children tend to derive scalar inferences far less often than adults do (e.g., Noveck 2001; Chierchia et al. 2001; Papafragou and Musolino 2003; Guasti et al. 2005; Foppolo et al. 2012, among many others). For example, Noveck (2001) presented underinformative sentences like (10) to a group of 8-year-old children, 10-year-old children, and adults. The participants were asked if they agreed with the sentences. Participants who derived the scalar inference in (11) were expected to disagree with the sentence.

(10) Some giraffes have long necks.

(11) Not all giraffes have long necks.

Noveck reported that child participants rejected the underinformative statements far less often than the adult participants did. This pattern of responses was taken as evidence that children derived fewer inferences than adults did. This result is representative of much of the subsequent work investigating children's performance on scalar inferences. It is worth noting, however, that many of the previous developmental studies have focused on a fairly restricted set of scalar inferences, namely those involving the scalar quantifiers 'some/all', the logical connectives 'or/and', and the modals 'may/must'. More recent work has turned to investigating inferences outside these lexical scales.

Adult-like behaviour

In the recent formal semantics/pragmatics literature, a wide range of inferences have received a scalar inference-based analysis (Levinson, 2000; Spector, 2007; Fox, 2007; Klinedinst, 2007; Chemla, 2009a; Thomas, 2012; Romoli, 2013; Chierchia, 2013; Crnič et al., 2015). Subsequent acquisition studies of these diverse inferences have produced some surprising results (Zhou et al., 2013; Tieu et al., 2014, 2016; Hochstein et al., 2016; Bill et al., 2016; Singh et al., 2016; Tieu et al., 2017), including the finding that some of these inferences are derived quite readily by young children.

One example of an inference that children have been reported to compute readily is the *free choice inference*, investigated by Tieu et al. (2016). Tieu et al. investigated Mandarin-speaking 4-year-old children's interpretations of disjunction under deontic modals, as well as English-speaking 5-year-old children's interpretations of free choice 'any' under deontic modals. Both sentence types give rise to free choice inferences, which have received a scalar inference analysis in the literature (Kratzer and Shimoyama, 2002; Alonso-Ovalle, 2005; Fox, 2007; Klinedinst, 2007; Chemla, 2009a).[7] For example, a sentence like (12) gives rise to the free choice inference in (13). On the scalar inference account, the free choice inference's derivation involves the alternatives in (14) and (15), which correspond to

[7]See Zimmermann (2000); Geurts (2005) and Barker (2010) for alternative analyses.

the disjuncts of the originally asserted sentence in (12).[8]

(12) Kung Fu Panda may push the green car or the orange car.

(13) Kung Fu Panda may push the green car and Kung Fu Panda may push the orange car.

(14) Kung Fu panda may push the green car.

(15) Kung Fu Panda may push the orange car.

Tieu et al. (2016) used a Truth Value Judgment Task (Crain and Thornton, 1998), in which the child participants were presented with free choice sentences like (12), in contexts that were inconsistent with the relevant free choice inference. That is, the character mentioned in the target sentence (e.g., *Kung Fu Panda*) was only permitted to perform one of the mentioned actions (e.g., *push a green car*). The study compared children's performance with free choice inferences and standard scalar inferences involving plain disjunctions (i.e. exclusivity inferences). Tieu et al. found that the child participants derived the exclusivity inference from test sentences with plain disjunction at a typically low rate (18%), whereas they derived free choice inferences from sentences that contained a deontic modal verb at a significantly higher rate (91%).[9]

As another example, children have been reported to interpret numeral terms as *exact* (Papafragou and Musolino, 2003; Barner and Bachrach, 2010; Huang et al., 2013). On certain accounts, such an interpretation involves the derivation of a scalar inference (Horn, 1972; Levinson, 2000); for instance, 'two' is interpreted as 'exactly two' through the negation of the stronger alternatives 'three', 'four', etc.

Turning to *ad-hoc* inferences, Stiller et al. (2015) presented 2-4-year-old children with sentence descriptions like (16). In the item associated with the sentence in (16), the participants' task was to identify which of three faces the sentence was describing. The three faces created an ad-hoc scale (i.e <*face with no glasses and no hat, face with glasses but no hat, face with glasses and hat*>). If participants

[8]Note that according to these scalar inference analyses, the scalar inference process is applied twice (i.e. recursively) to the alternatives in (14) and (15), in order to generate the free choice inference in (13). For more information about this process see Fox (2007).

[9]Tieu et al. (2016) report similarly high rates of free choice inferences from free choice English 'any', and Huang and Crain (2014) report similar rates of free choice inferences for Mandarin *renhe* 'any'.

derived the relevant ad-hoc inference in (17) for the target sentence description, they were expected to select the face with glasses and no hat.

(16) My friend has glasses.

(17) My friend does not have a hat.

Stiller et al. reported that 3-year-old children and 4-year-old children derived ad-hoc inferences like (17) approximately 75% of the time.

In yet another example of children's relative success, Hochstein et al. (2016) investigated 4–5-year-old children's derivation of ignorance inferences. These authors reported that 5-year-old children derived ignorance inferences at a rate of 70%, while they computed exclusivity inferences only 30-40% of the time.

Finally, children have also been reported to succeed in accessing *exhaustive* interpretations similar to those derived through the scalar inference process, when interpreting 'only'-sentences involving *context-dependent* scales (Barner et al., 2011). Barner et al. (2011) presented 4-year-old children with pictures depicting three objects/animals (e.g., *cat, cow, dog*), all of which were doing the same activity (e.g., *sleeping*). Children were asked one of two questions; (18) or (19).

(18) Are only some of the animals sleeping?

(19) Are only the cat and the cow sleeping?

The question in (18) is associated with context-independent alternatives, <*some...all*>, whereas the question in (19) is associated with context-dependent alternatives, <*cat, cow, dog*>). A negative answer to these questions was taken as evidence of participants accessing an *exhaustive* interpretation like (20).

(20) Only some, not all, of the animals are sleeping.

Barner et al. reported that children accessed the exhaustive interpretation in the context-independent condition 33% of the time, whereas they did so in the context-dependent condition 86% of the time.

Explaining children's variable behaviour

The majority of previous studies investigating children's abilities with scalar inferences have reported low rates of derivation in children. However, some recent studies have revealed a class of inferences that children successfully compute. This disparity in findings calls for an explanation for children's variable performance. The need for an explanation is especially pressing on unifying accounts that treat many of the relevant inferences in the same way (e.g., Fox (2007) and Chierchia (2013)).

Two main kinds of explanations for the developmental findings have been proposed. On one account, children's difficulties with scalar inferences are attributed to their limitations in *pragmatic knowledge* (Noveck, 2001; Katsos and Bishop, 2011; Skordos and Papafragou, 2016). This line is taken, for example, by Katsos and Bishop (2011), who propose that children are more *pragmatically tolerant* than adults. On this view, even when children derive scalar inferences, they are nevertheless more likely than adults to accept a target sentence, despite the fact that the inference makes the sentence false. Similarly, Skordos and Papafragou (2016) propose that children have a limited understanding of which alternative sentences are relevant in a given context (see Noveck 2001 for a similar idea). This limitation is expected to affect children's derivation of scalar inferences, since alternative sentences need to be perceived as 'relevant' before they can be negated.

Such *pragmatic* explanations are able to account for the typical observation of low rates of scalar inferences in children. However, it is unclear how such accounts can be extended to cases where children perform at adult-like rates in deriving inferences (e.g., free choice inferences). One would expect the proposed limitation ascribed to children (pragmatic tolerance or limited understanding of relevance) to have a uniform effect across different types of inferences, contra recent findings.

A second kind of explanation that attempts to account for children's selective success on scalar inferences is one we will broadly term the *Alternatives-based approach*. This approach attributes children's difficulties to a limitation in how they engage with the alternative sentences involved in the derivation of inferences. Some proposals in this camp make reference to the processing resources required to generate or retrieve alternatives (e.g., Reinhart's 2006 'Reference Set Hypothesis' or Tieu et al.'s 2016 'Restricted Alternatives Hypothesis'; see also Chierchia et al. 2001; Gualmini et al. 2001). If children do not have the processing resources (e.g., verbal working memory) required to generate, retrieve, or compare alternatives, then they will not be able to compute the target inferences.

Another variant of the Alternatives-based approach makes reference to limitations in lexical scale knowledge rather than to processing limitations. This perspective posits that children have a non-adult-like knowledge of the lexical scales from which many scalar inferences are derived (Barner and Bachrach, 2010; Barner et al., 2011). The derivation of scalar inferences such as the exclusivity inference, licensed by sentences with disjunction, requires the hearer to replace one scalar term in a sentence with one of its stronger scale-mates. Children who have limited knowledge of the lexical scale will be less likely to generate the alternative sentences required to compute the inference, and therefore will derive inferences at lower rates than adults do.

As with the pragmatic-based explanations, the Alternatives-based approach explains why children struggle with certain scalar inferences, especially those that require lexical replacement of scalar terms. In contrast to the pragmatic-based explanations, however, the Alternatives-based approach also provides a way to capture the cases where children readily derive inferences. That is, it is plausible that providing information relating to the relevant alternative sentences in the environmental or linguistic contexts could reduce the processing and scale knowledge requirements associated with the derivation of relevant inferences. And notably, many of the inferences that children succeed in computing involve alternative sentences that can be composed from linguistic material contained in the asserted sentences. Generating or retrieving such alternatives is less demanding on processing resources, and eliminates the need for children to have already established the relevant scale in their mental lexicon. Therefore, in cases like these, the Alternatives-based approach expects children to be able to overcome the proposed limitations and derive the relevant inferences readily.

The next section illustrates the predictions of the Alternatives-based approach by focusing on scalar inferences that involve alternatives that are contained within the asserted sentences. Specifically, we will focus on testing the predictions of the Alternatives-based approach using the distributive and conjunctive inferences that have been proposed to arise when a disjunction phrase is embedded under a universal quantifier.

2.1.2 Testing the Alternatives-based approach

The distributive inference

As noted earlier, a sentence like (21) is associated with the distributive inference in (22).

(21) Every elephant caught a big butterfly or a small butterfly.

(22) Some elephant caught a big butterfly and some elephant caught a small butterfly.

Crucially, as outlined in the present chapter's appendix (Section 2.4.2), the alternatives that are required to compute the distributive inference are not generated through lexical replacement. Rather, the alternatives are sentences in which the universal quantifier is combined with the individual disjuncts from the asserted sentence, i.e. the sentences in (23) and (24).

(23) Every elephant caught a big butterfly.

(24) Every elephant caught a small butterfly.

Since the alternatives are provided to them within the asserted sentences, children should be able to overcome the Alternatives-based approach's proposed limitations. As a result, according to the Alternatives-based approach, children are expected to be able to compute distributive inferences like (22) when responding to sentences like (21).

The conjunctive inference

Before turning to the experiment, we wish to note the presence of another potential inference that could arise from disjunctive sentences like (21). Singh et al. (2016) present findings suggesting that, unlike adults, children sometimes derive a *conjunctive inference* like (25) from disjunctive sentences like (21) (see Tieu et al. (2017) for corroborating evidence from French-speaking children and Japanese-speaking children).

(25) Every elephant caught both a big butterfly and a small butterfly.

Note that children deriving the conjunctive inference is not problematic for the Alternatives-based approach. This is because the conjunctive inference is proposed to be derived from alternatives that are presented as part of the asserted sentence, much like the alternatives that are required for the distributive inference (see this chapter's appendix (Section 2.4.3) for details). Therefore, if we observe that children sometimes derive conjunctive inferences, this will be taken as consistent with the Alternatives-based approach. This conclusion is justified by the same reasoning we applied to distributive inferences: children are expected to successfully compute inferences when the relevant alternatives can be formulated using linguistic material from the asserted sentences themselves, rather than retrieved from their mental lexicon.

Acknowledging the possibility that children will compute conjunctive inferences, then, our experiment investigates children's success in computing inference-based interpretations of the test sentences, including both distributive inferences and conjunctive inferences.[10] The finding that children accessed an inference-based interpretation of sentences like (21) at an adult-like rate would provide compelling evidence in support of the Alternatives-based approach.

2.2 Experiment

The present experiment was devised to determine the extent to which children derive distributive or conjunctive inferences from sentences in which disjunction is embedded under the universal quantifier. Both kinds of inferences rely on alternatives that can be formulated using linguistic material contained in the asserted sentences. Therefore, the experimental hypothesis, based on the Alternatives-based approach, was that the child participants would compute an inference-based interpretation at adult-like rates. This finding would be in contrast to the findings of much previous research on inferences that are derived through the retrieval of alternative lexical items from a language user's mental lexicon. In these studies, the child participants were found to derive inferences at significantly lower rates

[10]Note that the conjunctive inference asymmetrically entails the distributive inference. This makes it difficult to determine whether or not the child participants derive both inferences simultaneously. We discuss this issue further below.

than adult participants.

2.2.1 Method

Participants

Seventeen monolingual English-speaking adults (Macquarie University undergraduate students, all females) and 20 monolingual English-speaking children (4;01-5;08, M=4;05, 5 females, 15 males) participated in the experiment. The child participants had no history of language impairment or delay. The adults took part in the experiment for course credit, or for a payment of $15AUD. Children were recruited from several on-campus daycares at Macquarie University, and from a Macquarie University child research participant database. Informed consent was obtained from the adult participants, and from the parent/guardian of the child participants.

Procedure

The experiment used a version of the Truth Value Judgment Task (TVJT) (Crain and Thornton, 1998). Participants listened to a series of short stories, illustrated with photographs of toy props, on a laptop computer. After each story, a puppet appeared on the computer screen to utter the target sentence. Participants were asked to judge whether the puppet's sentence was right or wrong. If the participant said that the puppet was wrong, they were asked to provide a justification ("Why?" or "What really happened in the story?"). Adults were tested using the same materials, but were asked to provide written responses, including written justifications for *yes*- and *no*-responses.

Each child was tested individually, either in the lab or in a quiet room at their preschool. Adult participants were tested simultaneously in groups of up to three. The test procedures were split over two sessions conducted 7–9 days apart, each lasting approximately 15 minutes. This was done to separate the DISTRIBUTIVE and NON-DISTRIBUTIVE conditions, which we describe in detail below.

Materials

Each participant was presented with four test sentences in the Non-Distributive condition and four test sentences in the Distributive condition, as well as two clearly true and two clearly false control sentences. The control sentences included the universal quantifier 'every', without disjunction. In addition, four filler sentences were included, for a total of 12 trials. The sentences were presented by the puppet using pre-recorded videos.

Test conditions

Each item consisted of a single story. In each story, there were three characters and two sets of objects. In turn, each of the characters considered the two sets of objects in front of them; they then made a decision to perform a pre-designated action on an object from one of the sets, and then proceeded to perform that action. Once all the characters had complected the action, the experimenter asked the puppet to describe what the characters had done in the story. The puppet responded with the test sentence. To illustrate a typical item, the story in (26) was associated with the test sentence in (27).

(26) Example of Distributive target story
This is a story about three elephants. The elephants have come to a park to catch butterflies. There are big butterflies and small butterflies. The big butterflies are hard to catch, but the small butterflies are easy to catch. Let's see which butterflies the elephants decide to catch. The first elephant says: "I am not very fast, so I will catch a small butterfly." So she catches a small butterfly. The second elephant says: "I am very fast, so I will catch a big butterfly." So he catches a big butterfly. The third elephant says: "I am also very fast, so I will also catch a big butterfly." So he catches a big butterfly. And here are the elephants, with the butterflies that they caught [see Figure 2.1].
Experimenter: Now let's see if <Puppet's name> was paying attention. <Puppet's name>, what did the elephants do?

(27) **Puppet:** Every elephant caught a big butterfly or a small butterfly.

Figure 2.1: Final picture associated with the story in (26). Each elephant is holding the butterfly they caught.

In the DISTRIBUTIVE condition, participants who computed the literal meaning were expected to accept the target sentences, since every elephant caught a big butterfly or a small butterfly. Participants who computed the distributive inference were also expected to accept the target sentences, since at least one elephant caught a big butterfly and at least one elephant caught a small butterfly. However, participants who computed the conjunctive inference were expected to reject the target sentences, since none of the elephants caught both a big butterfly and a small butterfly.

In the NON-DISTRIBUTIVE condition, all of the characters acted upon the same kind of object (e.g., all of the elephants caught a big butterfly). As in the DISTRIBUTIVE condition, participants who computed the literal meaning were expected to accept the target sentences. Participants who computed the conjunctive inference were again expected to reject the target sentences, since none of the elephants caught both a big butterfly and a small butterfly. Participants who computed the distributive inference were also expected to reject the target sentences, since none of the elephants caught a small butterfly. A summary of the expected pattern of responses in the two conditions is presented in Table 2.1.

	Distributive	**Non-Distributive**
Literal interpretation	*accept*	*accept*
Distributive interpretation	*accept*	*reject*
Conjunctive interpretation	*reject*	*reject*

Table 2.1: Expected responses to the DISTRIBUTIVE and NON-DISTRIBUTIVE target sentences according to the three target interpretations.

Controls and fillers

In addition to the eight test items, participants also received four control items involving control sentences containing the universal quantifier 'every' but no disjunction. These control items were designed to ensure that participants understood the basic meaning of the universal quantifier. Participants were presented with two clearly true and two clearly false control items. For example, on one false control item, three horses have to decide which of two vehicles to drive, a car or a boat. Two of the horses end up driving a boat, and one drives a car. The puppet then utters the false control sentence in (28).

(28) Every horse drove a boat.

In addition to the targets and controls, participants also received four filler items, which could either be associated with a *yes*-response or a *no*-response. The experimenter determined which judgment to elicit from the participant, depending on the number of *yes*-responses and *no*-responses the participant had produced on previous trials. In this way, the experimenter ensured that each participant produced no more than two *yes*-responses or two *no*-responses in a row.

2.2.2 Results

Participants had to correctly answer at least 75% of the filler items and at least 75% of the control items in order to be included in the analysis. All participants passed this criterion, and no participants were excluded from the analysis.

The performance by child and adult participants in the test conditions is presented in Figure 2.2, plotted as the proportion of test sentence rejections in the Distributive and Non-Distributive conditions.

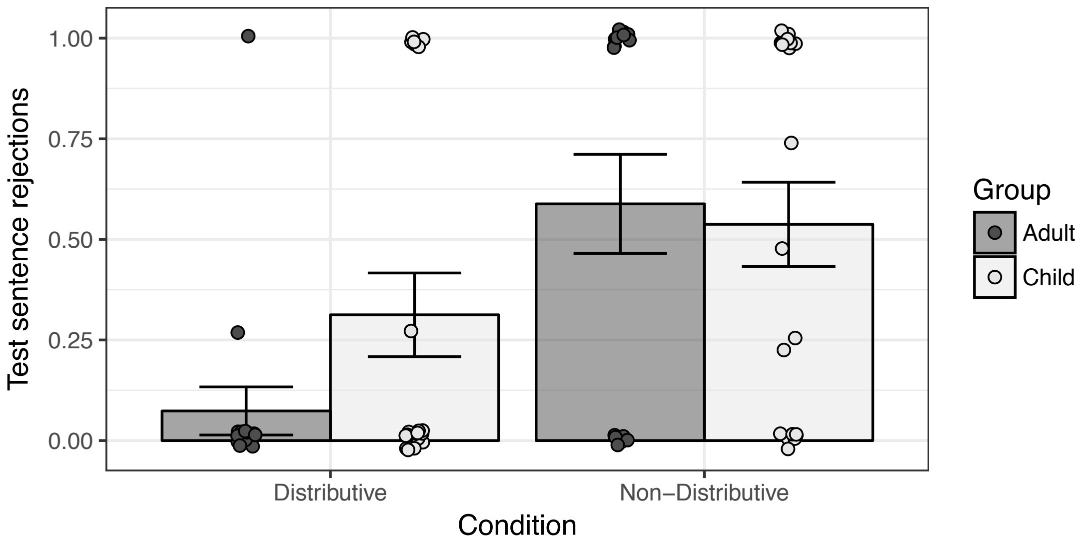

Figure 2.2: Proportion of test sentence rejections by Group and Condition. The vertical lines represent the standard error. Each dot represents an individual participant's mean rejection of test sentences in that condition. A horizontal jitter of .1 and vertical jitter of .025 were applied to the dots for easier visualisation.

The data were analysed using Generalized Mixed-Models (GLMER) in R (version 3.3.3) with the LanguageR package (version 1.4.1). In the analysis, we used a backward elimination procedure to compare the goodness-of-fit of the models. We started with the maximal structure that allowed the models to converge, following the recommendation by Barr et al. (2013). This maximal model contained two random intercepts (for participants and for items) and an individual adjustment of the condition for each participant (random slope). Group (Adult vs. Child), Condition (Distributive vs. Non-Distributive), and the interaction of Group-by-Condition were included as fixed effects. The reference categories in our analysis were 'Adult' for Group and 'Distributive' for Condition.

We evaluated whether each of the predictors (i.e. fixed effects) significantly contributed to the model's fit by comparing a model including that predictor against another that did not include it, using a χ-square test (Jaeger, 2008). Subsequently, we calculated the z-value, based on the Wald statistic, which allows for an estimation of the statistical significance of each predictor included in the model.

The fixed effect of Condition made a significant contribution to the fit of the model ($\chi^2(1) = 12.67, p < .001$), whereas the interaction of the two fixed effects ($\chi^2(1) = .38, p < .54$) and the fixed effect of Group ($\chi^2(1) = .01, p < .91$) did not contribute significantly to the fit of the model.

Estimated coefficients, their standard errors, Z-values, and associated p-values for the chosen model are reported in Table 2.2.

Predictor	**Est.**	**SE**	**Z Wald**	***p***
(Intercept)	10.74	2..29	4.70	<.001
Condition - Non-Distributive	-20.02	3.11	-6.43	<.001

Summary of fixed effects in the logit model (N = 296; log-likelihood = -68.2).

Table 2.2: The best fitting model of the data. Adult and DISTRIBUTIVE were used as reference levels (0) for Group and Condition factors, respectively.

The results reveal that test sentences in the NON-DISTRIBUTIVE condition were rejected at a significantly higher rate than test sentences in the DISTRIBUTIVE condition. No difference in test sentence rejection was found between the two groups.

To explore the response patterns of individual participants we categorised the participants into different groups based on the type of judgment each participant tended to produce in each of the test conditions. A summary of the categorisation criteria is presented in Table 2.3. For example, a participant was classified as *Distributive* if they accepted at least three of four test sentences in the DISTRIBUTIVE condition and rejected at least three of four test sentences in the NON-DISTRIBUTIVE condition.

Three children were excluded from the categorisation, as one gave mixed responses, and two produced a pattern of responses that was not predicted by any of the relevant interpretations. The remainder of the participants fell into one of the three categories. Table 2.4 presents the number of participants in each category.

Since our primary interest is in identifying whether children would derive an inference from the test sentences (whether distributive or conjunctive), Table 2.5

	Distributive	Non-Distributive
Literal group	*accept*	*accept*
Distributive group	*accept*	*reject*
Conjunctive group	*reject*	*reject*

Table 2.3: Expected response patterns to the DISTRIBUTIVE and NON-DISTRIBUTIVE test sentences, by different response groups.

	Child	Adult
Literal group	7	7
Distributive group	7	9
Conjunctive group	3	1

Table 2.4: Distribution of participant groups across the two response groups.

collapses the Distributive and Conjunctive group into a single *Inference* group; these are participants who computed one inference or the other.

As can be seen in Table 2.5, children and adults patterned identically in terms of their distributions in the literal group and inference group. Table 2.5 shows that children were as likely as adults to compute literal meanings, and as likely to compute an inference-based interpretation.

	Child	Adult
Literal group	7	7
Inference group	10	10

Table 2.5: Distribution of adults and children across the two response groups, collapsing Distributive and Conjunctive groups.

Justifications

In addition to *yes-/no*-responses, we also collected justifications from each participant. The justifications for the '*yes*-responses' tended to just repeat the basic elements of the relevant story. As a result, they do not provide any information regarding the interpretation the participant had accessed. The justifications for the *no*-responses were more informative, so we will focus on these.

First, we will consider the 83 justifications that were associated with participants' *no*-responses in the NON-DISTRIBUTIVE condition. This condition was inconsistent with the distributive inference as well as the conjunctive inference. These justifications were coded into 3 categories. The largest category contained 72% (60/83) of the justifications. The justifications in this category focused on the fact that all of the characters had acted on the one set of items. Here are three representative examples: *No, all the fairies ate a big cookie; No, they all did wash a big dish; No, all of them opened the big ones.* The next largest category, with 25% (21/83) of the justifications, were statements that focused even more explicitly on the fact that all of the characters had acted on the one set of items, with the additional information that that set was the only set of items the characters had acted on. Here are three representative justifications from this category: *No, everyone ate a big cookie only; No, the cows only washed the big dishes; No, every mouse only drank a big juice.* These two categories make up 97% of the participants' justifications. Both of these categories of justifications are consistent with the conclusion that the relevant *no*-response was a result of the participant having derived the distributive inference or the conjunctive inference, as the justifications focus on the fact that all the characters only acted on one kind of object. That is, they focus on the fact that these inferences are not upheld in the context. The final justification category for the NON-DISTRIBUTIVE condition contained only 2% (2/83) of the justifications. These were justifications that merely repeated what had happened in the story, or were unintelligible 'Other' responses.

Turning now to the 30 justifications associated with *no*-responses in the DISTRIBUTIVE condition. This condition was consistent with the distributive inference, but was not consistent with the conjunctive inference. These justifications were coded into 2 categories. The larger category contained 77% (23/30) of the justifications. These were statements that repeated what had happened in the story. Here are three representative statements of this kind: *No, two have big teas, and one got a little tea; No, one caught a small butterfly. Two caught a big butterfly; No, because two dogs drove a big car and one dog drove a little car.* These justifications do not reveal the exact semantic interpretation that led the participant to reject the test sentences. The other category contained 23% (7/30) of the justifications. Statements in this category focused on the fact that one of the characters bought something different than the other two characters bought. Here are representative statements: *No, one buyed a small one; No, Only one got a small tea; No, one girl ate a small one.* These justifications imply that the participant wanted all of the characters to act on the same kind of item (see Section 2.3.3 for more discussion of a possible reason underlying this kind of interpretation).

In sum, the *no*-response justifications in the NON-DISTRIBUTIVE condition were consistent with the conclusion that the participants had derived the distributive inference or the conjunctive inference. However, the *no*-response justifications in the DISTRIBUTIVE condition were less clear, with the bulk of participants merely repeating basic elements of the stories.

2.3 Discussion

The aim of this experiment was to investigate the predictions of the Alternatives-based approach (Chierchia et al., 2001; Gualmini et al., 2001; Reinhart, 2006; Barner and Bachrach, 2010; Barner et al., 2011; Tieu et al., 2016). Specifically, we aimed to test the prediction that children would be able to compute inferences that were derived from alternatives contained within the asserted sentences.

We investigated adults' and children's interpretations of asserted sentences like (29). Specifically, we were interested in the rates at which adults' and children's interpretations would include distributive inferences such as (30) and conjunctive inferences such as (31). Both these inferences are proposed to be derived as scalar inferences from alternatives that are presented as part of asserted sentences like (29) (Crnič et al., 2015; Singh et al., 2016).

(29) Every elephant caught a big butterfly or a small butterfly.

(30) Some elephant caught a big butterfly and some elephant caught a small butterfly.

(31) Every elephant caught both a big butterfly and a small butterfly.

The Alternatives-based approach expects children to readily access (i.e. at a similar rate as adults) an inference-based reading of assertions like (29), be it from deriving a distributive inference like (30) or from deriving a conjunctive inference like (31).

Our experimental results indicated that children indeed accessed inference-based readings of asserted sentences like (29), and importantly, did so as often as adults. This is revealed in the fact that there was no difference between the rates at which our participant groups rejected target sentences in the NON-DISTRIBUTIVE condition, which was inconsistent with both of the relevant inferences. This indicates that children were deriving inference-based interpretations as readily as adults, consistent with the predictions of the Alternatives-based approach.

In contrast, these results are not expected by explanations that appeal to more pragmatic mechanisms to account for children's behaviour. Our use of a binary-response truth value judgment task means that Katsos and Bishop (2011) would expect that children's 'pragmatic tolerance' would lead them to accept our target sentences more than adults do.

Moreover, it is not clear that the characteristics of the contexts were such that the alternatives were made particularly relevant. Therefore, Skordos and Papafragou (2016) and Noveck (2001) would also expect for children to not derive inference-based interpretations of our test sentences. The findings of the present study go against the expectations of these accounts, because the child participants rejected the test sentences at the same rate as adults. The findings support the conclusion that both groups derived an inference-based interpretation.

2.3.1 Quantifier spreading

As noted by an anonymous examiner, it is worth exploring whether our results could be attributed to children accessing a so called 'quantifier spreading' interpretation of our test sentences (Philip, 1991, 1995).

Previous work has found that children often reject sentences headed by a universal quantifier when they are used as descriptions of contexts where there are 'leftover' objects that are not matched with an agent. For example, it has been found that children will often reject a sentence like (32) when it is used to describe a context where there are some donkeys that are not being ridden (Philip, 1995).

(32) Every man is riding a donkey.

Turning to our experiment, in both of the test conditions there were objects that were not acted on (e.g., butterflies that were not caught). Therefore, if children were adopting this quantifier spreading interpretation, it should lead them to reject the target sentences in both of our test conditions. However, as the individual participant analysis reveals, there were only 3 children and 1 adult that provided such a response pattern (i.e. the conjunctive group). Therefore, while such an interpretation might provide an alternative explanation for this subset of participants, the bulk of participants' responses do not follow the response pattern expected by a quantifier spreading interpretation. For this reason, we do not take our conclusions regarding the nature of participants' interpretations to be significantly threatened by this possible alternative interpretation.

2.3.2 The conjunctive inference

The findings of the present experiment provide evidence that child participants derived inference-based interpretations at the same rate as adults. However, it is not clear that these inference-based interpretations provide evidence that any of the child participants consistently derived a conjunctive inference.

The proposal by Singh et al. (2016) suggests that at least some children, in contrast to adults, consistently derive conjunctive inferences from sentences with disjunction. If this is correct, then this kind of inference would be expected to be observed in the child participants' responses to the test sentences in the present experiment. In the present experiment, the outcomes in the Non-Distributive condition were inconsistent with both the conjunctive inference and with the distributive inference, so rejections of the test sentence in this condition could be used as evidence that the participants had derived one of these inferences or the other. In contrast, the Distributive condition was only consistent with the distributive inference; it was inconsistent with the conjunctive inference. So if a participant derived a conjunctive inference, they would be expected to reject the test sentences in this condition. Therefore, if children, in contrast to adults, derived conjunctive inferences, we would expect to find a difference between groups in the Distributive condition. However, we found no statistically significant difference between our participant groups in this condition. The absence of a between-group difference argues against the conclusion that children were deriving conjunctive inferences.

Having said that, an analysis of individual responses by participants revealed that 3 children and 1 adult produced a pattern of response patterns that might be associated with the conjunctive inference. These responses seem to suggest that at least some participants derived conjunctive inferences. However, as noted by an anonymous examiner, it is not clear how much we can conclude from the individual results, given this pattern is found in so few participants. That is, such response patterns could merely represent noise, rather than a genuine alternative interpretation. A certain level of skepticism regarding these interpretations is further justified by the fact that this response pattern was also generated by an adult, which, even according to Singh et al., is not expected.

Given these considerations, it is important to note that while our experiment produced reasonable evidence that children are deriving inference-based interpretations in an adult-like manner, the evidence that any of these interpretations were a result of having derived a conjunctive inference is less robust.

One possible reason why children were not deriving conjunctive inferences (or were deriving them at such a low rate) is that, while our experiment included a condition that was consistent with the distributive inference, there was no corresponding condition that was consistent with the conjunctive inference. Therefore, if children derived both inferences, the fact that the distributive inference was consistent with one of the experimental conditions may have weighed in its favour, leading participants to resolve the ambiguity by computing the distributive inference. To avoid creating this potential bias, future studies could include, in addition to our conditions, a condition in which the conjunctive inference is made true.

2.3.3 The egalitarian interpretation

Two of our child participants provided a response pattern that was not expected on any of the readings we identified. Instead, these participants appeared to access what could be called an *egalitarian* interpretation of the target sentences, meaning they would only accept the target sentence if each character had acted upon the same kind of object. That is, these two children rejected the target sentences in the Distributive condition and accepted them in the Non-Distributive condition. Moreover, the justifications these child participants offered for rejecting the target sentences in the Distributive condition indicated that they accepted the target sentences only if all of the characters performed the same action, e.g., *(No, because) only one got a small butterfly*; *(No, because) one drank a small tea, two drank a big tea.*

A similar pattern of responses was previously reported in a study by Boster and Crain (1993) and in a study by Kiguchi and Thornton (2016). Boster et al. tested 4-year-old children using sentences with disjunction, which were embedded under the universal quantifier 'every', as in the present study. The authors of both studies report that a subset of their child participants accepted the test sentences only if all of the relevant characters performed the same action. Boster et al. proposed that across languages, children initially adopt this egalitarian interpretation to ensure that other interpretations can be acquired on the basis of positive evidence. This proposal is based on the observation that the egalitarian interpretation makes sentences true in a narrow range of circumstances, making it the 'subset' interpretation. Speakers of a number of languages judge sentences to be true in circumstances corresponding to the egalitarian interpretation (e.g., Mandarin Chinese (An, 2015)).[11]

2.3.4 Conclusion

Recent developmental work has led to proposals that children's variable success in computing scalar inferences can be explained by appealing to the nature of the alternatives involved; we refer to this family of approaches as the *Alternatives-based approach.* This approach was the source of the specific prediction we investigated - namely that children, like adults, should derive inferences that rely on alternatives that are composed from linguistic material contained in the asserted sentence. To test this prediction, we used sentences in which disjunction was embedded under a universal quantifier, e.g., *every elephant caught a big butterfly or a small butterfly.* Such sentences give rise to the distributive inference that *some elephants caught a big butterfly and some elephants caught a small butterfly.* Crucially, this inference

[11] As noted by an anonymous examiner, despite the continuity that the presence of an egalitarian interpretation would have with these previous results, in the present experiment this pattern of responses was only found in a very small number of participants (2 children). Therefore, there is a good chance that this pattern of responses is a result of noise in the data, rather than a genuine interpretation strategy. Therefore, as with the conjunctive inference, we only claim to have tentative evidence that participants are genuinely accessing such an interpretation.

can be derived as a scalar inference using alternatives that can be formulated using the contents of the asserted sentence. The findings of the present study revealed that children derived the distributive inference at adult-like rates. In addition, three of the child participants computed the conjunctive inference that *every elephant caught both a big butterfly and a small butterfly.* This finding is consistent with results reported in Singh et al. (2016) and Tieu et al. (2017). Our findings of distributive and conjunctive inferences are in line with the predictions of the Alternatives-based approach, and provide support for it as a viable explanation of children's variable performance in computing scalar inferences.

2.4 Appendix

In this appendix, we outline how, from the assertion in (33), the exclusive inference in (34) can be derived as a scalar inference. Following this we will outline how the distributive inference and the conjunctive inference can be accounted for through this same scalar inference process.

(33) The elephant caught a big butterfly or a small butterfly.

(34) It's not the case that the elephant caught both a big butterfly and a small butterfly.

There is on-going debate regarding the exact mechanism underlying scalar inferences (Chierchia, 2006; Fox, 2007; Geurts, 2010). As the main contribution of this chapter does not hinge on assuming any one particular account, we can remain theory-neutral.[12]

2.4.1 Deriving the exclusivity inference

The asserted sentence in (35) typically gives rise to the exclusivity inference in (37), through the scalar inference process, as laid out in steps (39a)-(39c).

(35) **Asserted sentence:** The elephant caught a big butterfly or a small butterfly.

[12]The one place where this neutrality is not possible is in our description of the conjunctive inference in Section 2.4.3, which is analysed within the *Grammatical account* of scalar inference derivation (Chierchia, 2006; Fox, 2007).

(36) **Alternative sentences:**[13] The elephant caught a big butterfly and a small butterfly.

(37) **Negated alternative sentences:** It's not the case that the elephant caught a big butterfly and a small butterfly.

(38) **Asserted sentence + Negated alternative sentences:** The elephant caught a big butterfly or a small butterfly AND It's not the case that the elephant caught a big butterfly and a small butterfly.

(39) a. The asserted sentence in (35) is spoken, instead of the alternative sentences in (36).

b. The alternative sentences in (36) that are stronger than the asserted sentence in (35) are negated to generate the negated alternative sentences in (37).

c. If a negated alternative sentence in (37) does not contradict the asserted sentence in (35), it is combined with the asserted sentence in (35), thereby generating the final meaning in (38).

2.4.2 Deriving the distributive inference

When the scalar inference process is applied to a sentence like (40) and the asserted sentence's individual disjuncts are included in the set of alternative sentences, then, in addition to deriving the exclusivity inference in (41), the distributive inference in (42) is derived.

(40) Every elephant caught a big butterfly or a small butterfly.

(41) It's not the case that every elephant caught a big butterfly and a small butterfly.

(42) Some elephant caught a big butterfly and some elephant caught a small butterfly.

[13] As we will show in Section 2.4.2, the set alternative sentences for a disjunctive asserted sentence like (35) can also include the asserted sentence's individual disjuncts. However, as these disjuncts are not involved in the derivation of the exclusivity inference, for the sake of simplicity, we do not not include them in this outline.

The steps of through which the distributive inference is derived as a scalar inference from the asserted sentence in (43) are laid out in (47a)-(47c)

(43) **Asserted sentence:** Every elephant caught a big butterfly or a small butterfly.

(44) **Alternative sentences:** Every elephant caught a big butterfly and a small butterfly; Every elephant caught a big butterfly; Every elephant caught a small butterfly.

(45) **Negated alternative sentences:** It's not the case that every elephant caught a big butterfly and a small butterfly AND It's not the case that every elephant caught a big butterfly AND It's not the case that every elephant caught a small butterfly.

(46) **Asserted sentence + Negated alternative sentences:** Every elephant caught a big butterfly or a small butterfly AND It's not the case that every elephant caught a big butterfly and a small butterfly AND It's not the case that every elephant caught a big butterfly AND It's not the case that every elephant caught a small butterfly.

(47)
a. The asserted sentence in (43) is spoken, instead of the alternative sentences in (44).
b. The alternative sentences in (44) that are stronger than the asserted sentence in (43) are negated to generate the negated alternative sentences in (45).
c. If a negated alternative sentence in (45) does not contradict the asserted sentence in (43), it is combined with the asserted sentence in (43), thereby generating the final meaning in (46).

The final meaning generated through this process is quite complex and is composed of multiple parts, which we will now unpack. There is the *literal meaning* in (48), the *exclusivity inference* in (49), and the two negated individual disjuncts in (50a) and (50b), which, when considered in conjunction with the *literal meaning*, generate the *distributive inference* in (50c).

(48) **Literal meaning:** Each elephant is such that it caught a big butterfly or small butterfly.

(49) **Exclusivity inference:** It's not the case that every elephant caught both a big butterfly and a small butterfly.

(50) **Distributive inference:**

a. It's not the case that every elephant caught a big butterfly.

b. It's not the case that every elephant caught a small butterfly.

c. Some elephant caught a big butterfly and some elephant caught a small butterfly.

In this way, it is possible for the distributive inference to be analysed as a scalar inference, derived from an asserted sentence like (43).[14]

2.4.3 Deriving the conjunctive inference

Before we can outline how the *conjunctive inference* is proposed to be derived as a scalar inference, from an asserted sentence like (51), there are three things that we need to mention.

(51) Every elephant caught a big butterfly or a small butterfly.

First, the derivation of the conjunctive inference from a sentence like (51) is not possible on all accounts of scalar inferences. That is, in order to capture the conjunctive inference as a scalar inference, derived from a sentence like (51), the more recent *Grammatical account* of scalar inferences needs to be assumed. This is because the scalar inference process needs to be carried out at the *sentence-internally*, rather than at the *whole-sentence* level, a feat that is only possible on the Grammatical account (for more information on this account see Chierchia (2006); Fox (2007); Chierchia (2013, 2017)).

Second, it needs to be assumed that, in addition to its literal meaning, each alternative sentence also includes any inferences that would have been derived from it had it been the asserted sentence.[15] For example, if the individual disjuncts

[14]Note that Crnič et al. (2015) argue against deriving the distributive inference through this more traditional scalar inference process. However, the scalar inference process they do argue for is the same with respect to the nature of the alternative sentences. That is, the Alternatives-based approach makes the same predictions regarding children's behaviour. We therefore do not go into the details of their proposal here.

[15]In other words, the scalar inference process needs to be applied 'recursively'.

of the assertion in (51), i.e. (52) and (53), were asserted sentences themselves, language users would often derive the inference in (54), known as the *exhaustive inference*.

(52) Every elephant caught a big butterfly.

(53) Every elephant caught a small butterfly.

(54) And nothing else.

Including then, the individual disjuncts of the sentence in (51), (52), and (53), each with its exhaustive inference, generates the final 'enhanced' alternatives in (55) and (56).

(55) Every elephant caught only a big butterfly.

(56) Every elephant caught only a small butterfly.

Finally, it needs to be assumed that, unlike adults, children do not generate the conjunctive alternative sentence in (57).[16] In fact, this is why the conjunctive inference is proposed to be the only possible inference of disjunction for children; if the conjunctive alternative were available, conjunctive inference derivation would be blocked.[17]

(57) Every elephant caught a big butterfly and a small butterfly.

Singh et al. (2016) justify this assumption by proposing that children are limited in their generation of alternative sentences (along the lines proposed by the Alternatives-based approach); in particular, children are proposed to experience difficulties generating alternative sentences involving lexical replacement, such as the conjunctive alternative in (57).

With the inclusion then, of these three assumptions, the steps through which the conjunctive inference is derived through the scalar inference process, applied to an asserted sentence like (58), are presented in (62a)-(62c).

[16]See Singh et al. (2016) for further discussion of this point.

[17]See Davidson (2013) and Bowler (2014) for evidence of conjunctive inferences of disjunction in adult speakers of American Sign Language and Warlpiri, two languages that do not contain a conjunctive alternative.

(58) **Asserted sentence:** Every elephant caught a big butterfly or a small butterfly.

(59) **Alternative sentences:** Every elephant caught only a big butterfly; Every elephant caught only a small butterfly.

(60) **Negated alternative sentences:** It's not the case that every elephant caught only a big butterfly AND It's not the case that every elephant caught only a small butterfly.

(61) **Asserted sentence + Negated alternative sentences:** Every elephant caught a big butterfly or a small butterfly AND It's not the case that every elephant caught only a big butterfly AND It's not the case that every elephant caught only a small butterfly.

(62)
a. The asserted sentence in (58) is spoken, instead of the alternative sentences in (59).
b. The alternative sentences in (59) that are stronger than the asserted sentence in (58) are negated to generate the negated alternative sentences in (60).
c. If a negated alternative sentence in (60) does not contradict the asserted sentence in (58), it is combined with the asserted sentence in (58), thereby generating the final meaning in (61).

Again, the final meaning generated by this process is complex. It includes the *literal meaning* in (63) and the negated 'enhanced' alternative sentences in (64a) and (64b), which when considered in conjunction with the *literal meaning*, generate the *conjunctive inference* in (64c).

(63) **Literal meaning:** Each elephant is such that it caught a big butterfly or a small butterfly.

(64) **Conjunctive inference:**
a. It's not the case that every elephant caught only a big butterfly.
b. It's not the case that every elephant caught only a small butterfly.
c. Every elephant caught both a big butterfly and a small butterfly.

In this way it is possible to account for the derivation of the conjunctive inference from the asserted sentence in (58) through the scalar inference process.

Chapter 3

Children's free choice inferences

This chapter is based on the following paper:

Bill, C., Pagliarini, E., Romoli, J., Tieu, L., and Crain, S. (2017). Children's derivation of universal free choice inferences. In preparation.

Author contributions:[1]
All authors assisted in the conception and design of the experiment. **CB** and EP created the materials for the experiment, with feedback from SC. **CB** conducted the data collection, with assistance from EP. **CB** conducted the statistical analysis, assisted LT. All sections in the paper were drafted by **CB**. Paper revisions were contributed to by all authors by providing feedback, as well as by directly editing the paper.

[1]Authors referred to by their initials

Abstract

The initial wave of results from the relevant literature on child language found that children as old as 10 struggled to derive scalar inferences (Noveck 2001; Chierchia et al. 2001; Papafragou and Musolino 2003; Guasti et al. 2005, among others). However, as a wider variety of scalar inferences were investigated, it became clear that there were a handful of scalar inferences children readily computed (Papafragou and Musolino, 2003; Barner and Bachrach, 2010; Stiller et al., 2015; Tieu et al., 2016; Hochstein et al., 2016). Recent attempts to explain the observed variation in children's behaviour include a class of explanations that we will refer to as the *Alternatives-based approach.* This approach attributes children's variable success in deriving scalar inferences to limitations that affect children's ability to compose the alternative sentences involved. The proposal is that young children have limitations in processing or in scale knowledge that hinder their capacity to derive the alternative sentences required for the computation of certain scalar inferences. The Alternatives-based approach predicts that children will successfully derive a scalar inference if the alternatives can be composed using linguistic material that is contained in the asserted sentence. The current study evaluates this prediction by investigating children's behaviour with two kinds of free choice inference. The findings of the present study are consistent with the Alternatives-based approach. Children were found to derive the relevant free choice inferences at a higher rate than has been reported in studies of several other kinds of inferences, although children's level of inference derivation was still less than that of adults. We conclude that the results provide support for the Alternatives-based approach as a viable explanation of children's behaviour. Moreover, we discuss the extent to which the findings can help adjudicate between competing theories about the formal linguistic mechanisms used in the derivation of scalar inferences.

3.1 Introduction

In daily conversation, speakers and hearers make inferences that extend beyond the literal semantic content of the sentences produced. For instance, when presented with a sentence like (1), language users will often make the inference in (2). The inference in (2) is not entailed by the literal content of the sentence in (1). If (1) entailed (2), then the negation of (2) would be a contradiction. However, the sentence in (3) shows that the sentence in (1) can be amended in a manner that cancels the inference in (2), without generating a contradiction.

(1) The dog carried a green stone or a purple stone.

(2) The dog didn't carry both a green stone and a purple stone.

(3) The dog carried a green stone or a purple stone...in fact it carried both.

The fact that the sentence in (3) is not contradictory shows that (2) is an inference drawn from (1), rather than an entailment. This particular inference is known as an *exclusivity inference*, but it is just one of a large class of inferences called *scalar inferences*[2] (Horn, 1972; Grice, 1975; Gazdar, 1979; Levinson, 2000).

There is a traditional explanation for how scalar inferences are derived, including the exclusivity inference illustrated in (2). The traditional account is based on principles of general reasoning, in combination with social conventions that govern linguistic contributions of speakers and hearers in everyday conversation. These conventions have been articulated in a number of submaxims, including those in (4) and (5) (Horn, 1972; Grice, 1975).

(4) **First Maxim of Quantity:** Make your contribution as informative as is required (for the purposes of the exchange) (Grice, 1975).

(5) **Second Maxim of Quality:** Do not say that for which you lack evidence (Grice, 1975).

According to the traditional account of scalar inferences, the maxim in (4) dictates that if the speaker of a sentence with disjunction, as in (1), could have

[2]Also known as *scalar implicatures*.

produced the more informative sentence with conjunction, presented in (6), then she would have.[3]

(6) The dog carried a green stone and a purple stone.

The maxim in (5) dictates that, because the speaker did not select the more informative sentence, the hearer is entitled to infer that the speaker was not in possession of sufficient evidence to assert the more informative sentence in (6). Assuming that the hearer takes the speaker to be well-informed, the hearer infers that the speaker believes (6) to be false. The hearer therefore enters the negation of (6) into her mental model of the conversation context (see Sauerland (2004)). This explains the derivation of the exclusivity inference in (2).

Contemporary accounts of scalar inferences have all been influenced to some extent by this traditional account. One contemporary account that closely resembles the traditional account is known as the *Pragmatic account*, examples of which are presented in Sauerland (2004) and Geurts (2010). The Pragmatic account follows the traditional explanation by appealing to conversational norms and general principles of reasoning to explain how scalar inferences are derived. In contrast to the Pragmatic account, there is an alternative account known as the *Grammatical account*. The Grammatical account embeds the computation of scalar inferences into the recursive computational system for generating sentence interpretations. On this account, the derivation of scalar inferences invokes a covert grammatical operator, called 'EXH', which is shorthand for 'exhaustify' (e.g., Chierchia 2006; Fox 2007; Chierchia 2013, 2017). For our purposes, it is sufficient to define EXH as a command to apply the traditional scalar inference process to the linguistic material within its scope.[4]

The Pragmatic account and the Grammatical account do not always agree on which inferences should be classified as scalar inferences. One disagreement concerns the classification of sentences like (7). In (7), disjunction is combined with the deontic modal verb phrase 'is allowed to'. This combination is typically associated with a *free choice inference*, as indicated in (8).

[3]The strength of a sentence's informativity is based on entailment relations. That is, if sentence-A asymmetrically entails sentence-B, then sentence-A is more informative.

[4]For a more comprehensive and formal definition of EXH see Fox (2007). Also see Chierchia (2006) where the operator is referred to as a silent *Only* operator (shortened to *O*).

(7) The dog is allowed to carry a green stone or a purple stone.

(8) The dog is allowed to carry a green stone and the dog is allowed to carry a purple stone.

As with the exclusivity inference, the free choice inference in (8), licensed by a disjunctive statement, such as (7), can be nullified without the threat of a contradiction. For example, if the speaker chooses to add a continuation to (7) such as "...but I don't know which", as in (9), the free choice inference is cancelled, without leading to a contradiction.

(9) The dog is allowed to carry a green stone or a purple stone...but I don't know which.

For present purposes the main point is that, for reasons that we will detail later, certain types of free choice inference are able to be classified as scalar inferences by the Grammatical account, but not by the Pragmatic account. This leads to different empirical predictions by the alternative accounts. The Grammatical account predicts that language users engage the same inferential mechanisms (i.e. EXH) in processing all types of free choice inference as they do in processing other scalar inferences. Therefore, the Grammatical account expects for language users' behaviour with free choice inferences to mirror that of other scalar inferences. The Pragmatic account could accommodate this finding, but it is also consistent with a difference in language users' behaviour towards free choice inferences and other scalar inferences. Given this difference in the relative falsifiability of these accounts, we will take empirical results that are consistent with the Grammatical account's narrower range of expectations as providing support for it, over the Pragmatic account.

In addition to inspiring the identified theoretical analyses, scalar inferences have also been the focus of a significant amount of experimental work. This experimental research includes work on adult sentence processing, and on child language. The present thesis is concerned with scalar inferences in child language.

There have been numerous studies of children's knowledge of scalar inferences. The initial studies consistently reported that children struggled with scalar inferences, whereas adults experienced no corresponding difficulties (Noveck 2001; Chierchia et al. 2001; Papafragou and Musolino 2003; Guasti et al. 2005, among many others). Despite the bulk of initial reports following this pattern, some subsequent research, including our investigation in Chapter 2, has reported children's successful computation of a handful of scalar inferences (Papafragou and Musolino, 2003; Tieu et al., 2016; Hochstein et al., 2016). This variation in children's performance with scalar inferences calls for an explanation. One kind of explanation has been particularly adept at accounting for children's variable behaviour, as we will discuss. We will refer to this class of explanations as the *Alternatives-based approach*.

The Alternatives-based approach attributes children's variable behaviour with scalar inferences to limitations affecting their ability to form alternatives to the asserted sentences, which is a required step in scalar inference computation (e.g., Chierchia et al. 2001; Gualmini et al. 2001; Reinhart 2006; Barner and Bachrach 2010; Barner et al. 2011; Tieu et al. 2016). In many cases, scalar inferences must be derived using alternatives that contain lexical items accessed from the child's mental lexicon. The Alternatives-based approach contends that such inferences are more difficult for children to compute, due to the impositions they make on processing resources and scale knowledge. One prediction the Alternatives-based approach makes is that children will more easily derive a scalar inference, if the alternatives that are required for deriving the inference can be composed using constituents of the asserted sentence. One inference that fits this description is the free choice inference, illustrated in example (8). In this way, the Alternatives-based approach is able to account for children's insensitivity to some, but not all, scalar inferences.

The present chapter reports the findings of an experimental investigation of children's behaviour with two kinds of free choice inferences. To anticipate the results, children proved sensitive to both kinds of free choice inferences. The child participants derived these inferences at a higher rate than has been observed in much of the previous literature. It should be pointed out, however, that the child participants in the experiment did not achieve the same level of successful performance as the adult participants did. We discuss both children's relative success, and the observed differences between children's performance and that of adults, with reference to the theories we have introduced. We conclude, first, that the high rate at which children compute free choice inferences, although not adult-like, is in the direction that is anticipated by the Alternatives-based approach. We conclude, second, that children's high rate of success in deriving free choice inferences is in line with the Grammatical account, which anticipates that children will derive free choice inferences and other, equivalent, scalar inferences at similar rates.

The rest of the chapter proceeds as follows. First, we provide a more detailed description of the competing accounts of scalar inferences, with particular emphasis on the predictions these accounts make for children's behaviour. Then, we describe the previous developmental literature documenting children's variable success in computing different kinds of scalar inferences. The next topic is the Alternatives-based approach, and how this approach accounts for children's variable behaviour. We conclude that discussion by identifying a prediction of the approach for children's behaviour. Specifically that, according to the Alternatives-based approach, children are expected to be successful in computing free choice inferences, because the alternatives to the asserted sentence can be formulated using linguistic material drawn from the asserted sentence itself. Following this, we describe an experiment that was designed to test these predictions of the Alternatives-based approach. Next, we present the findings of the experiment, and conclude by discussing the extent to which the findings comply with the Alternatives-based approach. Finally, we remark on how the findings bear on the alternative accounts of scalar inferences.

3.1.1 Competing theories of scalar inference derivation

The Pragmatic account and the Grammatical account provide distinct explanations of how scalar inferences are derived. While the Pragmatic account follows the traditional account in basing scalar inference derivation on general reasoning and conversational norms, the Grammatical account postulates the application of a covert grammatical operator EXH. One way these mechanisms differ is in regards to how they can be applied to a sentence. The grammatical operator EXH can be applied at the level of the *whole sentence*, but it can also be *embedded* in the derivation, i.e., sentence internally. For example, EXH can be applied at two sites in sentence (10). It can be applied at the whole sentence level, as in (11). The application of the operator EXH at this level yields the inference in

(12). The Pragmatic account can also captures the inference in (12) as a scalar inference, as it is derived through the scalar inference process being applied at the whole-sentence level.

(10) The dog is allowed to carry a green stone or a purple stone.

(11) EXH[EXH[The dog is allowed to carry a green stone or a purple stone.]][5]

(12) The dog is allowed to carry a green stone and the dog is allowed to carry a purple stone.

On the Grammatical account, however, it is also possible for the EXH operator to be embedded within sentences. For example, in the case of a sentence like (13), EXH can be embedded under the universal quantifier, as in (14). This yields the inference in (15). This inference is not immediately expected by the Pragmatic account because the mechanisms invoked by the Pragmatic account only apply after the basic meaning of a sentence has been derived, i.e., at the level of the whole sentence. That is, unlike the Grammatical account, according to the Pragmatic account, it is not possible for the scalar inference process to be applied to only part of a sentence, as in (14).

(13) Every dog is allowed to carry a green stone or a purple stone.

(14) Every dog EXH[EXH[x is allowed to carry a green stone or a purple stone.]]

(15) Every dog is allowed to carry a green stone and every dog is allowed to carry a purple stone.

To summarise, both the Pragmatic account and the Grammatical account are able to classify inferences like (12) as scalar inferences. However, this is not the case for inferences like (15). The Pragmatic account does not claim that inferences like (15) are impossible to derive, but just that they are not able to be derived as a scalar inference. For example, Geurts and Pouscoulous (2009b) present an an alternative, non-scalar inference explanation of how the inference in (15) could be

[5]Note that, in order to derive the inference in (12) and (15), EXH needs to be applied recursively, as shown in (11) and (14). The reason why EXH needs to be applied in this way is presented in this chapter's appendix (Section 3.4.2 and Section 3.4.3). Both the Grammatical account (Fox, 2007) and the Pragmatic account (Spector, 2007) can apply the scalar inference process in this way.

derived from a sentence like (13). These differences mean that, while the Grammatical account expects language users' behaviour with the inference in (15) to mirror that of other scalar inferences, the Pragmatic account would be consistent with a wider range of outcomes. Adults tend to be successful in deriving inferences, regardless of whether they are derived through the scalar inference process or some other mechanism. Children's abilities, however, are more sporadic, and may provide a unique opportunity to adjudicate between these different accounts of scalar inference derivation. Before we present the experiment designed to take advantage of this opportunity, we will review the findings of previous studies investigating children's performance in deriving scalar inferences, and the attempts that have been offered to explain children's behaviour.

3.1.2 A review of scalar inferences in child language

Children's behaviour with scalar inferences has been the focus of a significant amount of experimental work for almost two decades. The bulk of this work has found that children derived scalar inferences at a lower rate than adults (Noveck 2001; Chierchia et al. 2001; Papafragou and Musolino 2003; Guasti et al. 2005; Foppolo et al. 2012, among others). For example, the second experiment in Chierchia et al. (2001) used a *Truth Value Judgment Task* (Crain and Thornton, 1998) to test 3-6 year old children's interpretations of sentences like (16). Specifically, Chierchia et al. wanted to investigate whether children would derive from such sentences, exclusivity inferences like (17).

(16) Every boy chose a skateboard or a bike.

(17) Not every boy chose both a skateboard and a bike.

Chierchia at al. presented test sentences like (16) as descriptions of contexts where the relevant exclusivity inference was false - for example, where *every boy chose both a skateboard and a bike.* As the relevant exclusivity inference in (17) was inconsistent with the context, rejection of target sentences by participants was interpreted as evidence that the exclusivity inference had been computed. The Chierchia et al. study found that the eight adult participants they tested computed the exclusivity inference 100% of the time. However, the fifteen child participants they tested only computed the exclusivity inference 50% of the time.

The majority of, particularly earlier, studies investigating children's behaviour with scalar inferences produced similar results to Chierchia et al. (2001), with children deriving the inferences at a low rate, as compared to adults (Noveck, 2001; Gualmini et al., 2001; Papafragou and Musolino, 2003; Guasti et al., 2005). However, a small number of studies have found children deriving a scalar inference at a high rate. One of the earliest instances of this behaviour was reported in a paper by Papafragou and Musolino (2003). Papafragou et al. investigated the rates at which children derived three different inferences proposed to be scalar inferences, including an inference associated with sentences like (18), namely, the *exactly-n inference* in (19).

(18) Two horses jumped over the fence.

(19) Exactly two horses jumped over the fence.

Papafragou et al. used a Truth Value Judgment Task (Crain and Thornton, 1998) to investigate the interpretations of 30 Greek-speaking 4-5 year olds. They also included a control group of 30 Greek-speaking adults. Papafragou et al. presented their test sentences as descriptions of contexts that were inconsistent with the relevant scalar inference. A rejection of a target sentence by a participant was interpreted as evidence that the participant had derived the relevant scalar inference. Papafragou et al. found that, while children derived all three inferences (including (19)) at a significantly lower rate than adults, they derived the exactly-n inference more readily than they did the other two scalar inferences. Similar results have been found in other studies investigating children's interpretation of numerals (Barner and Bachrach, 2010; Huang et al., 2013).

Children have also been found to readily derive *ad-hoc* inferences in a study by Stiller et al. (2015). Stiller et al. presented 2-4-year-old children with three faces and a target sentences like (20). The participant's task was to identify which of three faces the sentence was describing. The three faces created the ad-hoc scale in (22). Stiller et al. were interested in whether children would derive the relevant ad-hoc inference, such as the inference in (21) from sentences like (20).

(20) My friend has glasses

(21) My friend does not have a hat.

(22) <Face with no glasses and no hat, Face with glasses but no hat, Face with glasses and hat>

The selection of the picture described by the inference (e.g., face with glasses but no hat) by a participant was interpreted as evidence that the participant had derived an ad-hoc inference. Stiller et al. reported that 3-year-old children and 4-year-old children derived ad-hoc inferences approximately 75% of the time, which is substantially higher than the rates at which children have been found to derive more classical scalar inferences (e.g., Chierchia et al. (2001)).

Another example of children readily deriving a scalar inference is presented in a study by Hochstein et al. (2016). Hochstein et al. used a variant of the *Felicity Judgment Task* (Chierchia et al., 2001) to explore children's interpretations of sentences with disjunction like (23). Such sentences have been associated with *ignorance inferences* like (24) (Gazdar, 1979). Hochstein et al. presented participants with target sentences like (23) as descriptions of contexts like (25).

(23) The bear took a cup or a plate.

(24) The speaker is ignorant as to whether the bear took a cup and as to whether the bear took a plate.

(25) The bear takes the cup, but leaves the plate.

The participants' task was to choose which of two puppets had produced the target sentence; one puppet was blindfolded, but one could see. The selection of the blindfolded puppet by a participant was interpreted as evidence that the participant had derived an ignorance inference.

Hochstein et al. (2016) tested two groups of child participants. One group was comprised of 4-year-olds. The other group was comprised of 5-year-olds. There was no difference in the rate at which the group of 4-year-olds derived the ignorance inference versus the exclusivity inference (both ≈30%). However, while the 5-year-old group derived the exclusivity inference at about the same low rate (≈30%), they derived the ignorance inference at a significantly higher rate (≈70%).

Another study by Tieu et al. (2016) used a Truth Value Judgment Task (Crain and Thornton, 1998) to investigate Mandarin-speaking children's interpretations of sentences like (26). Tieu et al. were specifically interested in whether these children would derived the associated free choice inferences, like (27), from such sentences. There are several theoretical proposals suggesting that free choice inferences are derived as scalar inferences (Kratzer and Shimoyama, 2002; Alonso-Ovalle, 2005; Fox, 2007; Klinedinst, 2007; Chemla, 2009a).[6]

(26) Kung Fu Panda may push the green car or the orange car.

(27) Kung Fu Panda may push the green car and Kung Fu Panda may push the orange car.

The Mandarin-speaking child participants ranged in age from 3;07 to 4;09, with an average age of 4;03. The child participants were presented with the Mandarin word-by-word analogues of English sentences like (26). Mandarin interpretations of sentences like (26) were presented to children as descriptions of contexts that were inconsistent with the associated free choice inferences, such as (27). A rejection of the test sentence was taken as evidence that the free choice inference had been derived. Tieu et al. found that child participants derived free choice inferences at a significantly higher rate (91%) than the exclusivity inference (18%).

Finally, a study by Barner et al. (2011) found that children successfully computed *context-dependent exhaustive interpretations* of sentences containing *only*.

In sum, while the bulk of early results found that children struggled to derive scalar inferences, later work revealed that there are certain scalar inferences that children derive readily. This variation in children's success with scalar inference computation presents a challenge for explanations attempting to capture children's behaviour in this area. We will consider some of these explanations in the next section.

3.1.3 Explaining children's scalar inferences

A variety of explanations attempting to account for children's behaviour with scalar inferences have been proposed. We will partition these explanations into two types, based on the kind of limitation they attribute to children. One group of explanations attribute children's insensitivity to scalar inferences to a limitation

[6]See Zimmermann (2000), Geurts (2005), and Barker (2010) for alternative, non-scalar inference analyses of free choice inferences.

in *pragmatic knowledge* (Noveck, 2001; Katsos and Bishop, 2011; Skordos and Papafragou, 2016). For example, Katsos and Bishop (2011) propose that children are less likely than adults to reject a target sentence based on an inference derived from it being inconsistent in the relevant context. They refer to this as children being *pragmatically tolerant*. In support of this explanation Katsos et al. present experimental results showing that, when given the opportunity, children will often make use of an *intermediate* judgment option to evaluate test sentences associated with scalar inferences that are inconsistent with the target context.

From this result Katsos and Bishop (2011) conclude that previous results, which were based on a traditional, binary judgment measure and were interpreted as evidence of children struggling to compute scalar inferences, would be better interpreted as experimental artefacts. One difficulty this conclusion faces is that, most straightforwardly, the effect of this pragmatic tolerance should be applied equally to all scalar inferences. Therefore, it does not explain why there are certain scalar inferences (e.g., free choice inferences (Tieu et al., 2016)) that children have been found to derive readily, despite the relevant experiments employing a binary response measure.

Another explanation from the Pragmatics-based approach is presented by Skordos and Papafragou (2016). Skordos et al. propose that at-least some of children's variable behaviour with scalar inferences can be attributed to limitations in their understanding of how the context affects the *relevance* of prospective alternative sentences. This limitation is proposed to inhibit children's ability to identify the relevant alternatives for an asserted sentence. This, in turn, prevents them from deriving the associated scalar inferences.[7] In support of this proposal, Skordos et al. present a series of experiments in which they manipulate elements of the context to make the relevance of various alternative sentences more or less salient. The results of these experiments suggest that children derive scalar inferences at differing rates depending on how relevant children perceive the alternative sentences to be. One advantage of this explanation over that proposed by Katsos and Bishop (2011) is that, in addition to accounting for children's low rates of scalar inference derivation, the limitation proposed by Skordos et al. is expected to vary in how it affects children's computation rates, across different contexts. That is,

[7]Note that a less developed version of this idea was proposed in Noveck (2001).

for example, Skordos et al. expect that in certain contexts (i.e., when the relevance of alternative sentences is highly salient), children will have an easier time deriving relevant scalar inferences.

While Skordos and Papafragou (2016)'s proposed limitation captures some of the variability in children's success with scalar inferences, there are some cases that it does not easily capture. For example, Tieu et al. (2016) found that children derived free choice inferences much more readily than exclusivity inferences. However, it is not clear that there was any context-based difference between the relevant conditions that would have impacted the saliency of the different alternatives. The same seems to be true for many of the other experiments where children have been found to readily derive scalar inferences (e.g., Papafragou and Musolino (2003); Barner and Bachrach (2010); Hochstein et al. (2016)). Consequently, it would appear that, at best, the explanation proposed by Skordos and Papafragou (2016) is only able to account for part of children's variable success with scalar inferences.[8]

Before presenting the second group of explanations, we would like to note that, although the explanatory power of the first group of explanations is limited in the ways we have identified, there is one option open to them, through which they could increase this explanatory power. That is, these explanations could interpret children's success at deriving a given inference (e.g., free choice inferences) as evidence that a scalar inference analysis is inappropriate for that inference, and that, an alternative analysis should be adopted or created. While adopting such a strategy would allow these explanations to increase their ability to account for all of children's behaviour with the inferences in question, it is not 'cost-free'. Adopting such a strategy would not only be clearly post-hoc, but would also result in having to abandon the gains in parsimony achieved by capturing so many inferences through the one scalar inference mechanism. Therefore, before resigning ourselves to the adoption of such a strategy, it seems worthwhile to explore whether there might be an adequate explanation of children's variable behaviour with the relevant inferences that does not require the abandonment of scalar inference analyses for them. One promising attempt at such an explanation is presented by the sec-

[8]In fact, Skordos and Papafragou (2016) themselves suggest that their proposed limitation may only be one of several limitations affecting children's behaviour with scalar inferences.

ond group of explanations we will present, which we call the *Alternatives-based approach*.

3.1.4 The Alternatives-based approach

The explanations that make up the Alternatives-based approach attribute children's behaviour with scalar inference derivation to some limitation in how they interact with the alternative sentences involved (Chierchia et al., 2001; Gualmini et al., 2001; Reinhart, 2006; Barner and Bachrach, 2010; Barner et al., 2011; Tieu et al., 2016; Singh et al., 2016). There are at-least two distinct proposals within this approach. One is that children have limited processing capacity, which makes it difficult for children to generate alternative sentences and/or to compare those alternative sentences to an asserted sentence (Chierchia et al., 2001; Gualmini et al., 2001; Reinhart, 2006; Tieu et al., 2016). The other proposal is that children have limited knowledge of the abstract lexical scales involved in the generation of many alternative sentences (e.g., <some...all>) (Barner and Bachrach, 2010; Barner et al., 2011). If children are limited in either of these proposed ways, then their ability to derive certain scalar inferences is expected to be affected.

Specifically, the Alternatives-based approach expects that children will have the most difficulty deriving scalar inferences that involve alternatives derived through the retrieval of alternative lexical items from the mental lexicon (e.g., exclusivity inferences). Notably, many of the scalar inferences for which children's behaviour has been investigated involve such alternatives. In this way, the Alternatives-based approach is similar to the *pragmatics-based* explanations, in that, it is able to account for why children have often been found to struggle with the computation of scalar inferences.

In contrast to the pragmatics-based explanations, the Alternatives-based approach does not stop at capturing children's difficulties; it also provides a systematic account of why children have been found to successfully derive scalar inferences in certain cases, but not in others. According to the Alternatives-based approach, the proposed limitations that children experience are not expected to apply consistently across scalar inferences. That is, many of the scalar inferences that children have computed readily involve alternative sentences that can be formulated using material that is explicitly presented in the linguistic or environmental context. For example, the linguistic material from which the relevant alternative sentences are composed could be contained in the asserted sentence. One example of such an asserted sentence is the sentence in (28), which is associated with the ignorance inference in (29). The ignorance inference in (29) is derived through the negation of the alternative sentences (30) and (31) (Gazdar, 1979). Crucially, the linguistic material from which these alternative sentences are composed is presented as part of the assertion in (28).

(28) The dog carried a green stone or a purple stone.

(29) The speaker is ignorant as to whether the dog carried a green stone and as to whether the dog carried a purple stone.

(30) The dog carried a green stone.

(31) The dog carried a purple stone.

Presenting this linguistic material as part of the asserted sentence plausibly reduces the processing and scale knowledge demands associated with the formulation of the relevant alternative sentences, by eliminating the need for children to access an item from a lexical scale. As the processing and scale knowledge requirements for formulating these alternatives are reduced, the Alternatives-based approach no longer expects for children to experience difficulties with them. As a result, the Alternatives-based approach predicts that children can readily derive scalar inferences that involve alternatives that can be formulated from linguistic material contained in the asserted sentence, such as the ignorance inference in (29).

Similarly, in the study by Stiller et al. (2015), the alternatives and their scalar relationship are explicitly presented in the environmental context. Presenting such content in this way also plausibly reduces the processing and scale-knowledge requirements associated with composing the relevant alternative sentences. As a result, the Alternatives-based approach also expects for children to derive the relevant inferences readily.[9]

[9] In fact, Skordos and Papafragou (2016)'s finding that children's performance improved when the alternatives were made more salient in the environmental context could be interpreted along similar lines.

In this way, the Alternatives-based approach is able to account for children successful computation of many inferences including, ad-hoc inferences (Stiller et al., 2015), free choice inferences (Tieu et al., 2016), ignorance inferences (Hochstein et al., 2016), and distributive inferences (see Chapter 2). Compared to the Pragmatics-based explanations, then, the Alternatives-based approach is able to capture more of children's behaviour; both their failure to derive many scalar inferences, as well as why, in certain cases, they succeed. Moreover, it accounts for this behaviour in a way that retains a scalar inference analysis for many of the inferences in question.

We should note, however, that the Alternatives-based approach is not always able to account for children's behaviour in a way that retains a scalar inference analysis for the relevant inference. That is, if children successfully derive an inference, despite the relevant material not being contained in the linguistic or environmental contexts, then even the Alternatives-based approach may be forced to abandon the inference to an alternative, non-scalar inference analysis. For example, a study by Bill et al. (2016) investigated children's interpretation of sentences like (32), to see whether they would derive the associated inference in (33). This inference is traditionally analysed as a *presupposition*. However, Romoli (2012) proposes that it might be better analysed as a scalar inference, computed through the negation of the alternative sentence (34).

(32) The bear didn't win the race.

(33) The bear participated in the race.

(34) The bear didn't participate in the race.

Notice that, the linguistic material from which the alternative in (34) is composed, is not presented as a part of the asserted sentence in (32) (Romoli, 2012). Moreover Bill et al. did not present the relevant material in the environmental context. Therefore, according to the Alternatives-based approach, children should experience difficulty deriving the inference in (33). However, Bill et al. found that children derived the inference in (33) at a high rate (>75%). Therefore, in this case, even the Alternatives-based approach is unable to provide an account of children's behaviour that retains a scalar inference analysis for the inference in question. As a result, the Alternatives-based approach is forced to conclude, as Bill et al. do, that a non-scalar analysis is more appropriate for the inference in question. In one sense, this observation is a strength, rather than a weakness of the Alternatives-based approach, because it shows that the approach is not *so powerful* that it can retain a scalar inference analysis for all relevant inferences, regardless of how children behave.[10]

In sum, the Alternatives-based approach provides an explanation of children's behaviour with scalar inferences that; a) accounts in a systematic way for the variable nature of this behaviour, and b) retains a scalar inference analysis for many of the relevant inferences. Furthermore, the Alternatives-based approach provides clear predictions regarding how children will behave with many scalar inferences that have not yet been investigated. For example, if an inference is analysed as being derived from alternatives that can be formulated from content contained in the asserted sentence, then children are expected to derive it readily. One way to test the Alternatives-based approach is to investigate such an inference, to see if children's behaviour is in line with this prediction. In this chapter, we carry out such a test by investigating children's behaviour with two kinds of free choice inferences.

3.1.5 Basic and universal free choice inferences

As mentioned earlier, sentences like (35) are associated with free choice inferences (henceforth *basic free choice inferences*) like (36). Similarly, sentences like (37) are associated with *universal free choice inferences* like (38).

(35) The dog is allowed to carry a green stone or a purple stone.

(36) The dog is allowed to carry a green stone and the dog is allowed to carry a purple stone.

(37) Every dog is allowed to carry a green stone or a purple stone.

[10] Children's behaviour with the exactly-n inference associated with numerals appears to be another example of children readily deriving a scalar inference that the Alternatives-based approach expects them to struggle with (Papafragou and Musolino, 2003; Barner and Bachrach, 2010; Huang et al., 2013). However, see Barner and Bachrach (2010) for an attempt to account for this behaviour in a way that is consistent both with the Alternatives-based approach, and with a scalar inference analysis of the exactly-n inference.

(38) Every dog is such that it is allowed to carry a green stone and it is allowed to carry a purple stone.

In contrast to basic free choice inferences, universal free choice inferences are only proposed to be scalar inferences according to the Grammatical account of scalar inference derivation (see this chapter's appendix (Section 3.4.3) for more detail on this point) (Chemla, 2009b). That is, as the basic free choice inference is derived through the application of the scalar inference process at the whole-sentence level, both accounts can capture it as a scalar inference. On the other hand, the universal free choice inference is derived through the application of the scalar inference process embedded under the universal quantifier, as in (39).

(39) Every dog EXH[EXH[x is allowed to carry a green stone or a purple stone.)

As this feat is only possible according to the Grammatical acocunt, the Pragmatic account must capture the universal free choice inference through some other mechanism.[11]

Given that only the Grammatical account provides a scalar inference analysis for universal free choice inferences, children behaving towards them in the manner predicted by the Alternatives-based approach's predictions, would provide support for the Grammatical account. Put another way, the Alternative-based Approach's predictions only relate to children's behaviour with scalar inferences, therefore, if children's behaviour is in line with these predictions, it is consistent with the idea that this inference is derived as a scalar inference, a feat that is only possible according to the Grammatical account. Therefore, investigating children's behaviour with universal free choice inferences allows us not only to test the Alternative-based approach's predictions, but also to provide some contribution to the on-going debate between these different theories of scalar inferences.

Before presenting our experiment, we should note that, in addition to the identified free choice inferences, assertions like (35) and (37) are also associated with inferences that Fox (2007) calls *anti-conjunctive inferences.* Namely, the assertion in (35) is associated with the *basic anti-conjunctive inference* in (40), and the assertion in (37) is associated with the *universal anti-conjunctive inference* in (41).

[11] See Geurts and Pouscoulous (2009b) for a sketch of such an alternative mechanism.

(40) The dog is not allowed to carry both a green stone and a purple stone.

(41) Every dog is such that it is not allowed to carry both a green stone and a purple stone.

Note that the anti-conjunctive inferences are consistent with the free choice inferences, in that, a dog can have free choice regarding which stone it carries, while not being allowed to carry both stones. Put another way, the dog must ultimately choose only one of the two options.

Moreover, sentences like (37) are also associated with the the *individual free choice inference* in (42) and the *individual anti-conjunctive inference* in (43). These inferences are through to be derived by applying the scalar inference process to the whole of the sentence in (37), as shown in (44).

(42) At least one dog is such that it is allowed to carry a green stone and it is allowed to carry a purple stone.

(43) At least one dog is such that it is not allowed to carry both a green stone and a purple stone.

(44) EXH[Every dog is allowed to carry a green stone or a purple stone.

While interesting in their own right, these inferences are not the focus of our investigation.[12] Therefore, we do not test adults' and children's derivation of them. However, as they are potential interpretations of one of our test sentences, we take them into account when designing our experiment.

In sum, our experiment is designed to contribute to two current theoretical issues. Firstly, it investigates the Alternatives-based approach's expectation that children are able to derive both basic and universal free choice inferences. Secondly, it explores whether, as the Grammatical account contends, participants' behaviour with both these inferences will be similar, indicating a similar derivation process. Our experiment does this by comparing the rates at which children and adults derive basic free choice inferences and universal free choice inferences. The relevant alternatives are contained in the assertion for both these free choice inferences, so the Alternatives-based approach expects for children to be successful in deriving

[12]See this chapter's appendix (Section 3.4.3) for more information on the derivation of the individual free choice inference and the individual anti-conjunctive inference. See also Chemla (2009b).

both of them. Moreover, the Grammatical account derives both these types of free choice inference through the same process, meaning they expect for participants to derive them both at a similar rate.

We turn now to present the details of our experiment, in which we investigated children's behaviour with basic free choice inferences and universal free choice inferences.

3.2 Experiment

The primary aim of our experiment was to investigate the rates at which children derive basic free choice inferences and universal free choice inferences. We did this to test the Alternatives-based approach's prediction that children would be successful in deriving both these inferences.

3.2.1 Method

Participants

Seventeen monolingual English-speaking adults (Macquarie University undergraduate students, all females) and 20 monolingual English-speaking children (4;0-5;10, M = 4;06, 10 females, 10 males) participated in the experiment. The child participants had no history of language delay or impairment. The adults took part in the experiment for course credit, or for a payment of $15.00. Children were recruited from several on-campus daycares at Macquarie University, and from a Macquarie University child research participant database. Informed consent was obtained from the adult participants, and from the parent/guardian of the child participants.

Procedure

The experiment used the Truth Value Judgment Task (Crain and Thornton, 1998). This task involves two experimenters. One experimenter tells a series of stories to the participant, using pictures or toy props. The other experimenter plays the role of a puppet who watches the stories along with the participant. Following each story, the puppet is asked by the first experimenter to describe the events that took place in the story. The puppet's description of the events includes a target sentence. The experimenter then asks the participant whether the puppet's description was right, and the participant responds with their judgment (i.e. *yes* or *no*). If the participant rejects the puppet's sentence, the experimenter asks the participant to provide a justification (i.e. "Why is <puppet's name> wrong?" or "What really happened in the story?"). The procedure was the same for adults, except that they were asked to provide written justifications for both *yes* and *no* responses.

Each child was tested individually, either in a lab, or in a quiet room at their daycare. Adult participants were tested simultaneously in groups of 1-3. The experimental session took approximately 30 minutes.

Materials

The experiment included two test conditions, four control conditions, and a filler condition. The test conditions included a BASIC-FCI condition, for which there were three items, and a UNIVERSAL-FCI condition, for which there were four items.

Test conditions

The BASIC-FCI condition items included test sentences like (45), describing contexts that were consistent with the sentence's *literal meaning* (46), as well as the associated *basic anti-conjunctive inference* (47)[13], but that were not consistent with the associated basic free choice inference (48). Therefore, a rejection of the test sentence in the BASIC-FCI condition was interpreted as evidence that the basic free choice inference had been computed.

(45) The dog is allowed to carry a green stone or a purple stone.

(46) **Literal meaning:** The dog is allowed to carry a green stone or the dog is allowed to carry a purple stone.

[13] See this chapter's appendix (Section 4.4) for more details on the anti-conjunctive inference.

(47) **Basic anti-conjunctive inference:** The dog is not allowed to carry both a green stone and a purple stone.

(48) **Basic free choice inference:** The dog is allowed to carry a green stone and the dog is allowed to carry a purple stone.

The UNIVERSAL-FCI condition items included target sentences like (49), used to describe contexts that were consistent with the sentence's *literal meaning* (50), its associated *individual anti-conjunctive inference* (51), its associated *universal anti-conjunctive inference* (52), and its associated *individual free choice inference* (53).[14] Crucially, the contexts were not consistent with the universal free choice inference in (54). Therefore, a rejection of the test sentence in a UNIVERSAL-FCI item was interpreted as evidence that the universal free choice inference had been computed.

(49) Every dog is allowed to carry a green stone or a purple stone.

(50) **Literal meaning:** Every dog is such that it is allowed to carry a green stone or it is allowed to carry a purple stone.

(51) **Individual anti-conjunctive inference:** At least one dog is such that it is not allowed to carry both a green stone and a purple stone.

(52) **Universal anti-conjunctive inference:** Every dog is such that it is not allowed to carry both a green stone and a purple stone.

(53) **Individual free choice inference:** At least one dog is such that it is allowed to carry a green stone and it is allowed to carry a purple stone.

(54) **Universal free choice inference:** Every dog is such that it is allowed to carry a green stone and it is allowed to carry a purple stone.

An example UNIVERSAL-FCI item is presented in (55); this story would be associated with the test sentence in (56).

[14] As can be seen in the test item example (55), we do not satisfy the anti-conjunctive inferences as explicitly as we do the individual free choice inference. This is because some theorists suggest that the anti-conjunctive inferences are quite weak, and in some cases, not present at all (Simons, 2005; Fox, 2007). As a result, we felt it would be sufficient for the contexts to be merely tacitly consistent with these inferences, rather than explicitly endorsing them.

(55) **Universal-FCI example test item:**
This is a story about three dogs. These three dogs have come to the beach to carry stones. There are green stones and purples stones. Zebra here is the owner of the stones, so she is going to tell the dogs the rules that they have to obey. Zebra says to the first dog, "You dog, you are allowed to carry a green stone. You are also allowed to carry a purple stone. It is up to you. Here are the stones that you are allowed to carry." <Zebra gives a green stone and a purple stone to the first dog>. Zebra then turns to the second dog and says,"You dog, you are allowed to carry a green stone. You are also allowed to carry a purple stone. It is up to you. Here are the stones that you are allowed to carry." <Zebra gives a green stone and a purple stone to the second dog>. Zebra then turns to the third dog and says, "Now you dog, you are only allowed to carry a green stone. Here is the stone that you are allowed to carry." <Zebra gives a green stone to the third dog>. [see Figure 3.1
Experimenter: *Now <Puppet's name>, can you tell us something about the rules?*

(56) **Puppet:** Every dog is allowed to carry a green stone or a purple stone.[15]

Control and filler conditions

In addition to the test conditions, we also included four control conditions. The control conditions were designed to ensure participants understood the basic meanings of critical elements of the test sentences - namely, the universal quantifier 'every' and the deontic modal phrase 'allowed to'. For each of these words/phrases, there was a total of four items, two with control sentences expected to elicit *yes*-judgments, and two with control sentences expected to elicit *no*-judgments. For example, in one of the EVERY_TRUE condition items, expected to elicit a *yes*-judgment, there are three elephants deciding which of two bags, *big or small*, they should each buy. Ultimately, they all end up buying a *big bag*, and the puppet

[15]The ordering of the disjuncts was counterbalanced, so that sometimes the relevant (falsified) disjunct was presented first, and sometimes second.

Figure 3.1: Final scene of the item presented in (55).

provides the control sentence in (57).

(57) Every elephant bought a big bag.

Finally, we also included one Filler condition item, which had two possible target sentences. One of the target sentences was designed to elicit a *yes*-response, while the other target sentence was designed to elicit a *no*-response. The sentence that was ultimately used was chosen based on a participant's responses to previous trials, so that we could reduce the possibility that participants would accept or reject more than two target sentences in a row. For example, if a participant had rejected the two items preceding the Filler item, then the puppet would provide the filler sentence designed to elicit a *yes*-response.

Design

Participants were presented with all of the experimental items over the course of one session, approximately half-an-hour in length.

All of the universal-FCI and basic-FCI test conditions items were presented in a blocked design, and always in the same order, so that participants were always presented with all of the BASIC-FCI items before being presented with any of the UNIVERSAL-FCI items.[16]

Apart from the blocking of the UNIVERSAL-FCI and BASIC-FCI conditions, the rest of the item order was pseudo-randomised. An initial random order was created, and then was slightly modified to ensure that, provided participants responded as expected in control and filler items, they would not accept or reject more than two target sentences in a row. Once this first ordering had been established, a second version was created, with items presented in the opposite order (but with the same block ordering of the UNIVERSAL-FCI and BASIC-FCI condition items).

3.2.2 Results

To be included in the final analysis, participants needed to provide the expected response in at least 7 out of the 9 control/filler items. All of the participants met this criterion and were included in the final analysis.

We collected both judgment and justification data from participants. We will consider each of these in turn, starting with the judgments.

Judgments

For each test item participants provided either a *yes*-judgment or a *no*-judgment for the test sentence. A *no*-judgment was taken as evidence that the relevant free choice inference had been derived. As shown in Figure 3.2, adults derived basic free choice inferences 100% of the time, and universal free choice inferences 98% of the time. In contrast, children derived basic free choice inferences 70% of the time, and universal free choice inferences 66% of the time. Furthermore, children's average rates of derivation are based on a bi-modal distribution. That is, most child participants were consistent in either always deriving or always not deriving each free choice inference.

[16]The conditions were blocked in this way so that children could get used to engaging with the simpler BASIC-FCI items, before being presented with the more complicated UNIVERSAL-FCI items. It was thought that this would reduce the chances of children failing to derive universal free choice inferences merely due to being presented with such complex stories and sentences 'out of the blue', rather than because they were actually unable to derive this kind of inference.

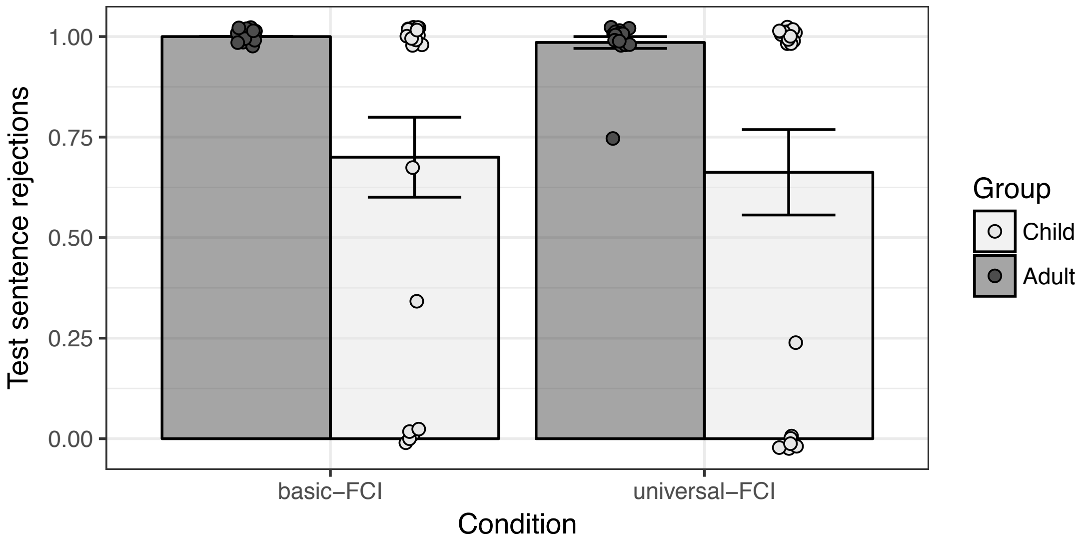

Figure 3.2: Mean proportion of test sentence rejections (i.e. inference-based responses) by Group and Condition. The vertical lines represent the standard error. Each dot represents an individual participant's mean rate of test sentence rejections in that condition. A horizontal jitter of .1 and a vertical jitter of .025 were applied to the dots for easier visualisation.

We analysed our data by first generating two values for each participant, corresponding to their mean proportion of target sentence rejections in each condition. We then ran a Pearson's Chi-square test to check for the possibility of an interaction, followed by a Wilcoxon rank-sum test to check for a main effect of group, and a Wilcoxon signed-rank test, to check for a main effect of condition. The Pearson's Chi-square test was not significant ($\chi^2(1) = 3.74$, $p = 1$), ruling out the possibility of an interaction effect, and the Wilcoxon signed-rank test found no significant effect of condition ($Z =$ -.63, $p =$ .53, $r =$ -.07). However, the Wilcoxon rank-sum test found a significant effect of group ($W = 232$, $p <$ .05, $r =$ -.41). Figure 4.2 indicates that the significant effect of group was driven by the fact that children derived both the basic free choice inference and the universal free choice inference at lower rates than adults did.

There was clearly a strong positive correlation between adults' rejection of test sentences in the BASIC-FCI condition and their rejection of test sentences in the UNIVERSAL-FCI condition, given their responses in both conditions are at ceiling. We conducted a Kendall's tau-b correlation test on children's responses. We found a strong and statistically significant positive correlation (Cohen, 1988) between children's rejection of test sentences in the BASIC-FCI condition and their rejection of test sentences in the UNIVERSAL-FCI condition ($\tau_b = .66$, $p < .01$).

Justifications

In addition to the truth value judgments, we also collected justifications from participants when they rejected a target sentence. We coded these justifications into different categories. In general, these justifications were consistent with participants having derived the relevant free choice inference.

Let us first consider the BASIC-FCI condition. In this condition, participants' rejections of the test sentences were accompanied by 93 justifications. These were coded into two categories. The larger category contained 82% (72/93) of the justifications. These justifications focused on the fact that the relevant character was only allowed to act on one of the objects as opposed to having the free choice to act on both mentioned objects. Representative examples include the following: *No, the dog can only use the yellow brush; No, the tiger is only allowed to buy the blue flower; No, just the green car.* These justifications are consistent with the conclusion that the participants had derived the basic free choice inference, as they focus on the fact that this inference is not upheld in the context. The other category of justifications in the BASIC-FCI condition contained the remaining 18% (17/93) of the justifications. These justifications either repeated basic elements of the story, or focused on irrelevant aspects of the story. The following are representative examples: *No, the dog got the yellow brush; No, otherwise the rhino stomps everywhere; No, the panda might steal the monkey's.* These justifications did not provide any insight as to the motivation for the relevant test sentence rejection.

Turning to the UNIVERSAL-FCI condition, the participants provided 120 justifications for rejecting test sentences in this condition. We coded these justifications into two categories. The larger category contained 89% (107/120) of these justifications. These focused on the fact that, unlike the other characters, one of the characters was only allowed to act on one kind of object. Examples include: *No, one dog is only allowed to carry a green stone; No, one horse is only allowed to* *buy a blue bead; No, only two are allowed white and red.* That is, these justifications invite us to conclude that the participants computed the universal free choice inference, which was not supported by the story contexts. The remaining 11% (13/120) of justifications in the UNIVERSAL-FCI condition tended to focus either on repeating basic elements of the story, or to point to irrelevant aspects of the story (e.g., *No, this cow got two flowers, and this cow got two flowers, and this cow got one flower; No, this one gets the green one; No, she first said the blue bead*). These justifications do not provide any special insights into the motivation for the associated test sentence rejection.

To summarise, the bulk of the justifications that participants provided for rejecting test sentences in both of the test conditions were consistent with the conclusion that they had derived the relevant free choice inference.

3.3 Discussion

The primary aim of our experiment was to investigate the predictions of the Alternatives-based approach. This approach consists of explanations that attribute children's variable success with scalar inferences to processing or scale knowledge limitations, which are proposed to affect their ability to generate the alternative sentences involved. Specifically, we tested the Alternatives-based approach's prediction that children will successfully derive a scalar inference if the alternatives can be composed from linguistic material contained in the asserted sentence. Two scalar inferences that fit this profile are the basic free choice inference and the universal free choice inference.

Our experiment found, within each participant group, no difference in the rates at which the two free choice inferences were derived. However, our experiment did find a difference between each participant group, in the rates at which each these free choice inferences were derived. We found that adults derived both free choice inferences more frequently than children. As we will now discuss, while the difference between adults and children requires some further explanation, the relatively high rate of free choice inference derivation is consistent with the Alternatives-based approach.

3.3.1 The Alternatives-based approach

Our finding that children did not derive either free choice inference at the same rate as adults is not in-line with the most straightforward interpretation of the Alternatives-based approach's predictions. That is, if 'ready derivation' is defined as 'at the same rate as adults', then our results would appear to not be completely in line with the Alternatives-based approach's prediction.

However, while the rate at which children derived free choice inferences in our experiment was not as high as adults, it was reasonably high (i.e. 65-70%). Moreover, this rate is substantially higher than the rate at which 4-5 year-old children have typically been found to derive scalar inferences derived through lexical replacement (50% or lower) (Noveck, 2001; Chierchia et al., 2001; Papafragou and Musolino, 2003; Guasti et al., 2005; Foppolo et al., 2012; Tieu et al., 2016; Hochstein et al., 2016). Therefore, it would appear that children's behaviour in our experiment was at least in the direction expected by the Alternatives-based approach. Nevertheless, the difference between the derivation rates of our participant groups calls for some further explanation.

One possibility is that this difference is a result of free choice inferences being derived through the recursive application of the scalar inference process. As we present in more detail in this chapter's appendix (Section 3.4.2), both of the free choice inferences we investigated are derived as scalar inferences through the recursive application of the scalar inference process to the relevant assertion, as shown in (58) and (59).

(58) EXH[EXH[The dog is allowed to carry a green stone or a purple stone.]]

(59) Every dog EXH[EXH[x is allowed to carry a green stone or a purple stone.]]

It is possible that this recursive application increases the processing requirements associated with deriving the basic free choice inference and the universal free choice inference. If this was so, it could explain why children in our experiment derived the target inferences at a lower rate than adults, even though the inferences were expected to be unaffected by the Alternatives-based approach's proposed limitations. To put it another way, the rate at which children derived the target free choice inferences was higher than has been found for more traditional scalar inferences derived through lexical replacement, as expected by the Alternatives-based approach. However, the processing costs associated with recursively applying the scalar inference process meant that, for these inferences, children were still not quite adult-like.

In order to gain further insight into this behaviour, future research would benefit from including a more 'classical' scalar inference (e.g., the exclusivity inference), as a direct comparison. Doing this would provide a more complete picture of how children's behaviour with free choice inferences compares to their behaviour with other kinds of scalar inferences.

3.3.2 Grammatical versus Pragmatic accounts of scalar inferences

The second aim of this chapter's experiment was to contribute to the on-going debate between the Grammatical and Pragmatic accounts of scalar inferences. As we mentioned, the Grammatical account and Pragmatic account explain scalar inference derivation by appealing to different mechanisms; conversational norms and general reasoning, versus the covert grammatical operator 'EXH'. Differences between these mechanisms result in these accounts having different perspectives on the inferences that are categorised as scalar inferences. That is, while both accounts can capture the basic free choice inference as a scalar inference[17], only the Grammatical approach captures the universal free choice inferences as a scalar inference (Fox, 2007). Therefore, the Grammatical account expects there to be a relationship between language users' behaviour with these two free choice inferences, whereas the Pragmatic account could accommodate a wider range of outcomes. Our finding, therefore, of a significant and strong positive correlation between participants' responses to the two types of free choice inference is consistent with the Grammatical account's proposal that they are derived through the same mechanism.

Moreover, the Grammatical account, but not the Pragmatic account, expects language users' behaviour with free choice inferences to mirror that of other equiv-

[17] Although, see Schulz (2006); Geurts and Pouscoulous (2009b) and Geurts (2010) for explanations of how free choice inferences could be derived through alternative, non-scalar inference mechanisms.

alent scalar inferences. As mentioned, according to the Alternatives-based approach, children should derive scalar inferences like the free choice inference readily. Therefore, to the extent that children's derivation of free choice inferences can be considered 'ready' (see Section 3.3.1 for discussion), their behaviour provides even more support for the Grammatical account's analysis of free choice inferences as scalar inferences.

3.3.3 Processing costs of the universal free choice inference

Within each participant group, we found no difference between the derivation rates of basic free choice inferences and universal free choice inferences. This is noteworthy because, while no theory we are aware of explicitly predicts there to be a difference, the sentences from which the universal free choice inferences are derived are clearly more complex than those of basic free choice inferences. Moreover, according to the Grammatical account, these free choice inferences differ in relation to the level at which the scalar inference process (i.e. EXH) is applied. That is, while EXH is thought to be applied at the whole sentence level for basic free choice inferences, it is embedded under the universal quantifier in the case of the universal free choice inference.

Given these differences, it would be unsurprising if their processing requirements differed. And, if such a disparity existed, it is possible it would have been reflected in participants' derivation rates. However, as noted, there was no difference in the rates at which these two free choice inferences were derived. While this result is worthy of note, it could just be that our measures are not sensitive enough to detect a difference in processing requirements between these inferences. Therefore, in order to explore this possibility further, future work could employ more sensitive measure (e.g., reaction time, eye-tracking).

3.3.4 Conclusion

We set out to test the expectations of the Alternatives-based approach; a theory that attributes children's variable behaviour with scalar inferences to certain limitations in how they interact with the alternative sentences involved. We did this by investigating children's behaviour with basic free choice inferences and universal free choice inferences to see whether, as the Alternatives-based approach expects, children would derive them both readily. We found that, while children derived both free choice inferences at a higher rate than has been typically found with scalar inferences derived through lexical replacement, the rate was still significantly lower than that of adults. This behaviour provides tentative support for the Alternatives-based approach, however, the difference in derivation rates between children adults requires further explanation.

Our experiment also contributed to the on-going debate between competing theories of scalar inference derivation. That is, while both accounts can capture the basic free choice inference as a scalar inference, only the Grammatical account does so for the universal free choice inference. Therefore, the strong positive correlation between participants' responses to the two types of free choice inference is consistent with the Grammatical account's suggestion that they are derived through the same process. Moreover, the fact that children's behaviour was in the direction expected by the Alternatives-based approach provides further support for the Grammatical account.

3.4 Appendix

In this Appendix, we show how the scalar inference process has been used to account for the derivation of basic and universal free choice inferences. Due to the fact that only the Grammatical account provides a consistent scalar inference-analysis for all free choice inferences, we will assume this approach in the following analyses. Note, however, that those inferences derived through application of the scalar inference process at the whole-sentence level (e.g., basic free choice inferences) can also be captured as scalar inferences by the Pragmatic account.

3.4.1 Deriving the exclusivity inference

To start we show how the process of computing scalar inferences has been used in the derivation of a traditional scalar inference; the exclusivity inference associated with disjunctive sentences. According to the Grammatical account, scalar inferences are derived as a result of the grammatical operator 'EXH' being applied to all or part of an asserted sentence. Based on this principle, the exclusivity inference is derived from asserted sentences like (60), through the steps outlined in (64a)-(64c).

(60) **Asserted sentence:** EXH[The dog carried a green stone or a purple stone.]

(61) **Alternative sentences:**[18] The dog carried a green stone and a purple stone.

(62) **Negated alternative sentences:** It's not the case that the dog carried a green stone and a purple stone.

(63) **Asserted sentence + Negated alternative sentences:** The dog carried a green stone or a purple stone AND It's not the case that the dog carried a green stone and a purple stone.

(64) a. The asserted sentence in (60) is spoken, instead of the alternative sentences in (61).

b. The alternative sentences in (61) that are stronger than the asserted sentence in (60) are negated to generate the negated alternative sentences in (62).

c. If a negated alternative sentence in (62) does not contradict the asserted sentence in (60), it is combined with the asserted sentence in (60), thereby creating the final meaning in (63).

We now turn to how this scalar inference process has been used to account for the derivation of free choice inferences.

3.4.2 Deriving the basic free choice inference

Applying the scalar inference process to a sentence like (65) results in the derivation of free choice inferences like (66). However, this case is different to exclusivity

[18]Note that, as we have already made some mention of, and as outlined in detail in Section 3.4.2, some versions of this process (e.g., Gazdar (1979)) also include the asserted sentence's individual disjuncts in the set of alternative sentences. However, as these alternative sentences are not actually required to derive the exclusivity inference, and as we present them in greater detail momentarily, we have left them out of this derivation.

inference in that the scalar inference process needs to be applied recursively, as shown in (67).

(65) The dog is allowed to carry a green stone or a purple stone.

(66) The dog is allowed to carry a green stone and the dog is allowed to carry a purple stone.

(67) EXH[EXH[The dog is allowed to carry a green stone or a purple stone.]]

Recursive application of the scalar inference process can be understood as meaning that each of the alternative sentence also includes any scalar inferences they would have communicated, had they been uttered as asserted sentences themselves (i.e. if EXH was applied to them) (Fox, 2007).[19] For example, the asserted sentence in (65) is thought to be associated with the conjunctive alternative sentence in (68) and the disjunct-alternative sentences in (69) and (70). For reasons presented by Fox (2007), which we will not go into here, the conjunctive alternative (68) is not able to be 'pre-exhaustified'. However, EXH is able to be applied to the disjunct-alternative sentences, which in each case results in the derivation of the inference in (71).

(68) The dog is allowed to carry a green stone and a purple stone.

(69) EXH[The dog is allowed to carry a green stone.]

(70) EXH[The dog is allowed to carry a purple stone.]

(71) The dog is not allowed to carry anything else.

Therefore, (69) and (70) are each combined with the inference in (71), which, results in alternatives that can be paraphrased as (72) and (73). As a result, the final set of alternative sentences associated with the assertion in (65) is comprised of the conjunctive alternative in (68), and the two pre-exhaustified disjunct-alternatives (72) and (73).

(72) The dog is only allowed to carry a green stone.

(73) The dog is only allowed to carry a purple stone.

[19] For a more comprehensive and formal definition of recursive exhaustification see Spector (2003), Chierchia (2006), and Fox (2007).

Applying then, the scalar inference process recursively to an asserted sentence like (68), as in (74), results in the derivation of a free choice inference through the steps outlined in (78a)-(78c).

(74) **Asserted sentence:** EXH[EXH[The dog is allowed to carry a green stone or a purple stone.]]

(75) **Alternative sentences:**[20] The dog is allowed to carry a green stone and a purple stone; The dog is only allowed to carry a green stone; The dog is only allowed to carry a purple stone.

(76) **Negated alternative sentences:** It's not the case that the dog is allowed to carry a green stone and a purple stone; It's not the case that the dog is only allowed to carry a green stone AND It's not the case the dog is only allowed to carry a purple stone.

(77) **Asserted sentence + Negated alternative sentences:** The dog is allowed to carry a green stone or a purple stone AND It's not the case that the dog is allowed to carry both a green stone and a purple stone AND It's not the case that the dog is only allowed to carry a green stone AND It's not the case that the dog is only allowed to carry a purple stone.

(78) a. The asserted sentence in (74) is spoken, instead of the alternative sentences in (75).

b. The alternative sentences in (75) that are stronger than the asserted sentence in (74) are negated to generate the negated alternative sentences in (76).

c. If a negated alternative sentence in (76) does not contradict the asserted sentence in (74), it is combined with the asserted sentence in (74), thereby creating the final meaning in (77).

The final meaning (77) is quite complex as it is comprised of multiple elements, which we will now unpack. First, there is the *literal meaning* in (79), next, there is the *basic anti-conjunctive inference* in (80), and finally, the negated disjunct-alternatives in (81), which, when considered in conjunction with the literal meaning

[20] Note that the disjunct alternatives included in this set have already had the scalar inference process applied to them once. That is, they have been 'pre-exhaustified'.

in (79), generate the *basic free choice inference* in (81a). Note that, the basic anti-conjunctive inference in (80) is consistent with the basic free choice inference in (81a), in that, it is possible that the dog is **allowed** to carry a green stone and the dog is **allowed** to carry a purple stone, while it also being the case that the dog is **not allowed**, ultimately, to carry both (i.e. the dog needs to make a choice regarding which stone it will carry).

(79) **Literal meaning:** The dog is allowed to carry a green stone or the dog is allowed to carry a purple stone.

(80) **Basic anti-conjunctive inference:** The dog is not allowed to carry both a green stone and a purple stone.

(81) **Basic free choice inference:** It's not the case that the dog is only allowed to carry a green stone AND It's not the case that the dog is only allowed to carry a purple stone.

   a. The dog is allowed to carry a green stone and the dog is allowed to carry a purple stone.

In sum, by recursively applying the scalar inference process it is possible to derive the basic free choice inference through the scalar inference process. Next, we will show how universal free choice inferences can be accounted for through this same process.

3.4.3 Deriving the individual and universal free choice inferences

Universal free choice inferences are also able to be derived through the scalar inference process. Although, in this case, universal free choice inferences are only derivable as scalar inferences according to the Grammatical account of scalar inference derivation. This is because their derivation involves the application of the scalar inference process sentence internally. That is, for a sentence like (82), a universal free choice inference like (83) is derived as a result of EXH being embedded under the universal quantifier, as in (84). Applying the scalar inference process in this way is only possible on the Grammatical account of scalar inference derivation.

(82) Every dog is allowed to carry a green stone or a purple stone.

(83) Every dog is such that it is allowed to carry a green stone and it is allowed to carry a purple stone.

(84) Every dog EXH[EXH[x is allowed to carry a green stone or a purple stone.]]

Before we show how the application of EXH at this embedded level results in the derivation of a universal free choice inference, we will present the meaning that is derived from the application of EXH at the whole-sentence level, as in (85).

(85) EXH[EXH[Every dog is allowed to carry a green stone or a purple stone.]]

When EXH is recursively applied to a sentence like (82) at the whole-sentence level, as in (85), the scalar inference process (outlined in (90a)-(90c)) results in the final meaning in (89).

(86) **Asserted sentence:** EXH[EXH[Every dog is allowed to carry a green stone or a purple stone.]]

(87) **Alternative sentences:** Every dog is allowed to carry a green stone and a purple stone; Every dog is only allowed to carry a green stone; Every dog is only allowed to carry a purple stone.

(88) **Negated alternative sentences:** It's not the case that every dog is allowed to carry a green stone and a purple stone AND It's not the case that every dog is only allowed to carry a green stone AND It's not the case that every dog is only allowed to carry a purple stone.

(89) **Asserted sentence + Negated alternative sentences:** Every dog is allowed to carry a green stone or a purple stone AND It is not the case that every dog is allowed to carry a green stone and a purple stone AND It is not the case that every dog is only allowed to carry a green stone AND It is not the case that the dog is only allowed to carry a purple stone.

(90)
a. The asserted sentence in (86) is spoken, instead of the alternative sentences in (87).
b. The alternative sentences in (87) that are stronger than the asserted sentence in (86) are negated to generate the negated alternative sentences in (88).

c. If a negated alternative sentence in (88) does not contradict the asserted sentence in (86), it is combined with the asserted sentence in (86), thereby creating the final meaning in (89).

Breaking down the final meaning in (89) into its composite parts; there is the *literal meaning* in (91), the *individual anti-conjunctive inference* in (92), and the negated disjunct-alternatives in (93), which, when considered in conjunction with the literal meaning in (91), generate the *individual free choice inference* in (93a). We use these labels because in this case, and in comparison with their *universal* counterparts, the inferences are *individual*, in that they are only required to be applicable to one entity (e.g., dog) in the relevant set to be satisfied.

(91) **Literal meaning:** Every dog is such that it is allowed to carry a green stone or it is allowed to carry a purple stone.

(92) **Individual anti-conjunctive inference:** At least one dog is such that it is not allowed to carry both a green stone and a purple stone.

(93) **Individual free choice inference:** It is not the case that every dog is only allowed to carry a green stone AND It is not the case that every dog is only allowed to carry a purple stone.

a. At least one dog is such that it is allowed to carry a green stone and it is allowed to carry a purple stone.

Turning now to the derivation of the universal free choice inference in. As mentioned, this is achieved by applying the scalar inference process sentence internally, that is, embedded under the universal quantifier, as in (84). Notice that the linguistic material in the scope of EXH in (84) is essentially the same as for the basic free choice inference (see (67)). Therefore, the specific steps associated with the application of the scalar inference process to this linguistic material are equivalent to those we presented for the basic free choice inference in Section 3.4.2. However, the final meaning is different, due to the influence of the universal quantifier 'every'. That is, the relevant inferences are *universal*, in that they are required to be applicable to every entity (e.g., dog) in the relevant set to be satisfied. Breaking down the final meaning into its composite parts, there is the *literal meaning* (94), the *universal anti-conjunctive inference* (95), and the negated disjunct-alternatives in (96), which, when considered in conjunction with the literal meaning in (94), generates the *universal free choice inference* in (96a).

(94) **Literal meaning:** Every dog is such that it is allowed to carry a green stone or it is allowed to carry a purple stone.

(95) **Universal anti-conjunctive inference:** Every dog is such that it is not allowed to carry both a green stone and a purple stone.

(96) **Universal free choice inference:** Every dog is such that it is not the case that it is only allowed to carry a purple stone AND Every dog is such that it is not the case that it is only allowed to carry a green stone.

   a. Every dog is such that it is allowed to carry a green stone and it is allowed to carry a purple stone.

In summary, it is possible to capture the basic free choice inference and the individual free choice inference through the recursive application of the scalar inference process at the whole-sentence level. Moreover, by embedding this process under the universal quantifier, a feat that is only possible according to the Grammatical account, the universal free choice inference can also be derived as a scalar inference.

Chapter 4

Children's inferences from sentences with *every...some*

This chapter is based on the following paper:

Bill, C., Pagliarini, E., Romoli, J., Tieu, L., and Crain, S. (2017). Children's interpretation of sentences with multiple scalar terms. In preparation.

Author contributions:[1]
All authors assisted in the conception and design of the experiment. **CB** and EP created the materials for the experiment, with feedback from SC. **CB** conducted the data collection, with assistance from EP. **CB** conducted the statistical analysis, assisted LT. All sections in the paper were drafted by **CB**. Paper revisions were contributed to by all authors by providing feedback, as well as by directly editing the paper.

[1]Authors referred to by their initials

Abstract

Previous research has revealed considerable variability in children's performance computing different types of scalar inferences (Noveck, 2001; Guasti et al., 2005; Barner and Bachrach, 2010; Tieu et al., 2016). One promising strategy to account for this variation is presented by the explanations comprising the *Alternatives-based approach* (Reinhart, 2006; Barner et al., 2011; Tieu et al., 2016). This approach proposes certain processing and scale knowledge limitations that reduce children's ability to derive the alternative sentences involved in scalar inference derivation. Moreover, the Alternatives-based approach suggests these limitations can be overcome by presenting the linguistic material, from which the alternative sentences are formulated, within the asserted sentence. The present chapter investigates one such sentence, which includes an existential expression, *some*, in the scope of the universal expression, *every*. The results of this experiment provide further support for the Alternatives-based approach, as well as insights into several other important theoretical issues.

4.1 Introduction

The interpretation of the sentence in (1) often includes the sentence in (2) as a part of its final meaning. Although it may seem otherwise, (2) is not a part of the literal meaning of (1). Rather, (2) is an inference licensed by the sentence in (1). Its status as an inference is clear from the fact that (2) can be explicitly negated without contradiction. One way to negate the inference in (2) is illustrated in (3). Although the continuation in (3) negates the inference in (2), there is no outright contradiction. The present chapter investigates children's knowledge of inferences like (2) - aspects of meaning that are inferred, but not part of the literal meaning of a sentence.

(1) The pig carried some of his rocks.

(2) The pig didn't carry all of his rocks.

(3) The pig carried some of his rocks...in fact he carried all of them

The inference from (1) to (2) is called a *scalar inference*.[2] The traditional account of how such inferences are derived was proposed in Grice (1975) and in Horn (1972). This account supposes that scalar inferences are based on a combination of general reasoning, norms of conversation, and knowledge of lexical scales. By a lexical scale, we mean scales of multiple, semantically related lexical items, ordered by informational strength, such as; <*some, many, most, all*>, <*or, and*>, and <*may, must*>.

The traditional account of scalar inference derivation proposes that they are derived in the following way. Suppose someone engaged in a conversation is presented with a sentence that contains a lexical item selected from one of these lexical scales, e.g., the lexical item 'some' in (1). We will call the person who utters the assertion in (1), the speaker, and the person who is presented with the assertion in (1), the hearer. In such circumstances, the hearer may reason as follows. If the speaker is aware of the facts, and is being cooperative, then the sentence in (1) conveys the speaker's intended meaning in the most perspicuous way possible. If an alternative sentence would have conveyed the speaker's intended meaning more directly, in virtue of containing a different lexical item on the relevant scale (say,

[2] Also known as a *scalar implicature*.

'all'), then the speaker would have used the alternative, as in (4). The fact that the speaker did not choose to utter the more informative sentence invites the hearer to infer that the speaker is likely to believe that the alternative sentence is false. To keep the hearer's mental model of the conversational context in close alignment with that of the speaker, the hearer enters the negation of the alternative into her mental model of the conversational context. The sentence in (1) causes the hearer to compute the inference in (2).

(4) The pig carried all of his rocks.

Contemporary accounts of scalar inferences derivation are all influenced in some way by this traditional account. Two of these accounts were introduced in previous chapters. These are the *Pragmatic account* and the *Grammatical account* (see, e.g., Section 3.1.1). The Pragmatic account follows the traditional account in supposing that scalar inferences are derived using general principles of reasoning and conversational norms (Spector, 2003; Sauerland, 2004; Geurts, 2010; Geurts and van Tiel, 2013). In contrast, the *Grammatical account* diverges from the traditional account. According to the Grammatical account, the derivation of scalar inferences requires a covert exhaustification operator (abbreviated 'EXH', short for 'exhaustify' (Fox, 2007)). As a result, the Grammatical account contends that scalar inferences are derived during the on-line compositional process in which sentence meanings are computed on the basis of their constituent parts. (Recanati, 2003; Chierchia, 2004, 2006; Fox, 2007; Chierchia et al., 2011; Chierchia, 2013). For our purposes the grammatical operator 'EXH' can be thought of as a command for the language user to carry out the traditional scalar inference process on the linguistic content in its scope.[3]

Because the Pragmatic and Grammatical accounts invoke different mechanisms, these accounts often differ in how they categorise certain inferences. For example, both accounts associate *EverySome* sentences like (5) with *NotEvery inferences* like (6) and with *None inferences* like (7). However, these inferences are analysed differently by these competing accounts.

(5) Every pig carried some of his rocks.

[3]For a more formal and detailed account of the 'EXH' operator see Fox (2007) or see Chierchia (2006), where it is called the silent 'only' operator.

(6) Not every pig carried all of his rocks.

(7) None of the pigs carried all of his rocks.

On the Grammatical account, each of these inferences are derived through the application of the EXH operator to different parts of the sentence in (5), so both inferences are analysed as being derived through the same scalar inference process. In contrast, the Pragmatic account only straightforwardly categorises the NotEvery inference in (6) as a scalar inference. To license the None inference in (7), researchers who adopt the Pragmatic account appeal either to a more complex instantiation of the scalar inference process (Spector, 2003; Sauerland, 2004; Chemla, 2009a), or to an alternative, non-scalar inference mechanism (Geurts, 2010). In these ways, the Pragmatic account is able to license the derivations of both the NotEvery inference and the None inference, but there is some difference between the inferential processes that are adopted for each inference. In contrast, the Grammatical account takes both inferences to be derived through the same process. As a result, the Grammatical account expects greater continuity between language users' behaviour with the two inferences. While the Pragmatic account could accommodate such an outcome, it could also accommodate a difference in language users' behaviour. Therefore, the Pragmatic account is less directly falsifiable.

In addition to theoretical work on scalar inferences, the past two decades have witnessed a great deal of experimental investigations of children's derivation of scalar inferences. The findings of these investigations have revealed a substantial variability in children's success in computing scalar inferences of different kinds. The bulk of the initial work was on 'classic' cases of scalar inferences. The findings suggested that children struggled to compute scalar inferences, especially when compared with adults (Noveck, 2001; Chierchia et al., 2001; Papafragou and Musolino, 2003; Guasti et al., 2005; Foppolo et al., 2012). However, as research has expanded to include a wider range of scalar inferences, it was discovered that children readily computed a handful of scalar inferences (Papafragou and Musolino, 2003; Barner et al., 2011; Tieu et al., 2016; Hochstein et al., 2016)[4]. This variability in behaviour presents a challenge for explanations attempting to capture

[4]See also Chapter 2 and Chapter 3 of the current thesis.

children's behaviour with scalar inferences.

One promising group of such explanations is called the *Alternatives-based approach*. The Alternatives-based approach stands out in its ability to successfully capture children's struggles in computing some, but not all, scalar inferences. The Alternatives-based approach attributes this variable behaviour to certain limitations affecting children's capacity to compose the alternative sentences from which many scalar inferences are derived (Barner and Bachrach, 2010; Barner et al., 2011; Tieu et al., 2016; Singh et al., 2016). When children fail to compute inferences, the Alternatives-based approach contends that it is because composing the relevant alternatives exceeds children's abilities. When children succeed in computing scalar inference, the Alternatives-based approach proposes that it is because certain aspects of the linguistic or environmental context have enabled children to overcome these limitations. For example, children are predicted to readily derive scalar inferences when the linguistic material that is required to formulate the relevant alternative sentences is made explicit in the asserted sentence. Considering the scalar inferences that children have been found to derive readily, many of them are associated with such asserted sentences. In sum, the Alternatives-based approach is a very successful account of children's variable behaviour with scalar inferences; both their successes and their failures.

In addition to accounting for children's previously reported behaviour, the Alternatives-based approach also makes predictions regarding children's behaviour with many scalar inferences that are yet to be investigated. As just noted, the Alternatives-based approach expects that children will succeed in deriving a scalar inference if the linguistic material involved in formulating the relevant alternative sentences can be gleaned from the asserted sentence. As we will outline in more detail later, two inferences that fit this criteria are the NotEvery inference in (6) and the None inference in (7). Therefore, the Alternatives-based approach predicts that children will successfully derive at least one of these inferences from EverySome sentences like (5).

Again, while the Alternatives-based approach expects children to access an inference-based interpretation of EverySome sentences like (5), it remains agnostic about which of the possible inferences in (6) or (7) this interpretation will be based on. There is a theory, however, that is not so ambivalent as to which inference children will make in response to EverySome sentences. Even more interestingly, this theory predicts that children and adults will prefer different inferences. The theory, advanced in Crain et al. (1994), proposes that when engaging in sentence interpretation adults and children are guided by different considerations. Children are guided by learnability considerations. Adopting a version of the *subset principle* (Berwick, 1985), Crain et al. propose that children will initially assign the interpretation that includes the informationally stronger inference - the None inference in (7). This initial assignment allows children to add other interpretations through exposure to positive evidence. In contrast, it is proposed that adults are guided by parsing considerations that will lead them to prefer the interpretation including the weaker inference - the NotEvery inference in (6). In sum, according to Crain et al., whereas children are expected to prefer the stronger None inference, in response to EverySome sentences, adults are expected to exhibit a preference for the weaker NotEvery inference.

In this chapter, we present an experimental investigation measuring the rates at which adults and children derive the NotEvery inference in (6) and the None inference in (7) from EverySome sentences like (5). The results of this experiment suggest that adults and children access a similar rate of inference-based interpretations. This result is consistent with the predictions of the Alternatives-based approach. Moreover, we find that the specific inference driving this interpretation differs between our groups. Children were found to access interpretations based on deriving the None inference, while adults' interpretations were a result of deriving the NotEvery inference. These results are in line with the Crain et al. (1994)'s predictions, and are also consistent with the Grammatical account of scalar inferences.

The rest of this chapter is structured as follows. First, we review the previous experimental literature on children's computation of scalar inferences, followed by an outline of the different explanations that have been proposed for children's behaviour. Then, we outline the predictions of the Alternatives-based approach, as well as the learnability and parsing considerations proposed in Crain et al. (1994), as they relate to the assignment of interpretations to EverySome sentences by adults and children. Next, we outline in more detail the two identified accounts of scalar inferences and the predictions they make for language users' interpretations of EverySome sentences. The details of the experiment, designed to test these predictions are presented next. We conclude by discussing how the findings bear on the Alternatives-based approach, Crain et al.'s proposal, and the identified accounts of scalar inferences.

4.1.1 Children's scalar inferences

Over the past few decades, a great deal of research has focused on children's ability to compute scalar inferences. The first wave of these investigations consistently produced results suggesting that children struggled to derive scalar inferences, as compared to adults (Noveck 2001; Papafragou and Musolino 2003; Huang and Snedeker 2009; Foppolo et al. 2012, among others). For example, Noveck (2001) used a *reasoning scenario* to investigate children's behaviour with a number of different scalar inferences. Participants were presented with sentences like (8) and asked whether they agreed with them or not. Note that, based on world knowledge the associated *OnlySome* scalar inference in (9) is false. Therefore, if a participant had derived the scalar inference in (9), they were expected to reject the test sentence in (8).

(8) Some giraffes have long necks.

(9) Not all giraffes have long necks.

Noveck (2001) ran the study with 31 8-year-olds, 30 10-year-olds, and 15 adults. All participants were native French speakers and the test sentences were presented in French. Noveck's results suggested that adults had derived OnlySome inferences like (9) 69% of the time, while the two child groups derived it 11%(8 y/o)-15%(10 y/o) of the time. This result, in conjunction with similar results from two other experiments presented in this paper lead Noveck to conclude that children are less likely than adults to derive scalar inferences. The bulk of results produced by experimental studies over the proceeding decade produced similar results.

While the studies done by Noveck (2001) and others displayed convincing evidence that children struggled to derive the target inferences, the studies were largely focused on a small group of scalar inferences (primarily those associated with the scales; 'some/all', 'or/and', and, 'might/must'). As these investigations continued, and particularly as the studies started to include a wider range of scalar inferences, a different pattern has started to emerge - namely, it has been found that there are a handful of inferences that children have been found to derive readily, including those investigated in Chapter 2 and Chapter 3 of the present thesis (Papafragou and Musolino, 2003; Barner and Bachrach, 2010; Tieu et al., 2016; Hochstein et al., 2016). For example, a study by Tieu et al. (2016) investigated children's interpretations of sentences like (10), to see whether they included *free choice inferences* like (11), an inference which has received a series of scalar inference analyses (Kratzer and Shimoyama, 2002; Alonso-Ovalle, 2005; Fox, 2007; Klinedinst, 2007; Chemla, 2009a).[5]

(10) Kung Fu Panda may push the green car or the orange car.

(11) Kung Fu Panda may push the green car and Kung Fu Panda may push the orange car.

Tieu et al. used a *Truth Value Judgment Task* (Crain and Thornton, 1998) and conducted their experiment with 22 Mandarin-speaking children. Tieu et al. presented participants with Mandarin translations of target sentences like (10) as descriptions of contexts in which the relevant free choice inference was false (e.g., where Kung Fu Panda may only push the orange car). Given this context, a rejection of the test sentence by a participant was interpreted as evidence that the participant had derived a free choice inference. Tieu et al. found that while children derived a more traditional scalar inference (the *exclusivity inference* from disjunctive sentences) at a typically low rate (18%), they derived free choice inferences, like (11), at a much higher rate (91%).

Similar results were produced in a study by Hochstein et al. (2016), which was investigating children's interpretations of sentences like (12), to see if they would include associated *ignorance inferences* like (13). That is, Hochstein et al. found evidence of 5-year-old children deriving the ignorance inference at a much higher rate (≈76%) than the exclusivity inference (≈30%).

(12) The bear took a cup or a plate.

[5] Although see Zimmermann (2000); Geurts (2005) and Barker (2010) for alternative analyses.

(13) The speaker is ignorant as to whether the bear took a cup and as to whether the bear took a plate.

Moreover, a study conducted by Stiller et al. (2015) tested whether children would derive *ad-hoc inferences* from sentences like (14). Specifically, Stiller et al. presented sentences like (14) to 2-to-4-year old children. Children were directed to identify which, of three faces, the test sentence was describing. The characteristics of the three faces resulted in the following ad-hoc scale, <face with no glasses and no hat, face with glasses but no hat, face with glasses and hat>. It was expected that if children derived the target ad-hoc inference in (15), they would select the face with glasses but no hat. Stiller et al. reported that 3- and 4-year old children derived such ad-hoc inferences readily, at a rate of approximately 75%.

(14) My friend has glasses.

(15) My friend does not have a hat.

A series of studies have also reported evidence of children readily deriving an *exactly* interpretation of numerals (i.e. *one* = *exactly one*) (Papafragou and Musolino, 2003; Barner and Bachrach, 2010; Huang et al., 2013), a meaning that is also proposed to be computed as a scalar inference. Finally, a study by Barner et al. (2011) found evidence of children accessing *context-based exhaustive interpretations* of certain sentences, an interpretation that is thought to be accessed through a similar process to scalar inferences.

In sum, while earlier studies found that children struggled to derive the target scalar inferences, more recently, there have been a small but growing number of studies that have found children deriving certain scalar inferences readily.

4.1.2 Explaining children's variable success in computing scalar inferences

A number of explanations have been proposed to account for children's variable success in computing scalar inferences. These explanations can be categorised into two different groups based on the nature of the limitation they attribute to children. We will consider each of these groups in turn.

The Pragmatics-based approach

First, there are explanations that attribute children's behaviour with scalar inferences to limitations in their *pragmatic knowledge* (Noveck, 2001; Katsos and Bishop, 2011; Skordos and Papafragou, 2016). For example, Katsos and Bishop (2011) propose that children are or more *pragmatically tolerant* than adults. This pragmatic tolerance is proposed to mean that, even if children were to derive a scalar inference, they would be less likely than adults to reject a target sentence, based on the inference making the sentence false. In support of this proposal Katsos et al. present results from an experiment in which children are asked to judge the 'acceptability' of a target sentence using a 3-point scale, rather than the traditional binary, true/false response options. Children in this experiment tended to use the intermediate judgment option when judging target sentences in which the literal content was true, but the scalar inference was false. Katsos et al. interpret this result as consistent with the idea that children's previously documented poor performance in scalar inference computation is a result of the response scale used, rather than a result of not having computed the scalar inference at all.

Similarly, Skordos and Papafragou (2016) propose that children have a limited understanding of which alternative sentences are relevant in a given context (see Noveck (2001) for a similar idea). Skordos et al. support this proposal with experimental results showing that, when the relevance of alternatives are made more salient, children derive scalar inferences more readily.

Before presenting the second group of explanations we would like to note that, while these Pragmatics-based explanations can account for why children have been seen often to struggle with scalar inference derivation, they are less successful in accounting for many of the cases in which children have been found to derive certain scalar inferences readily. This is because, the limitations attributed to children by this approach are, most straightforwardly, expected to affect children's derivation of scalar inferences consistently. As a result, this approach struggles to explain why, for example, Tieu et al. (2016) found children deriving free choice inferences significantly more than exclusivity inferences, despite the contexts being, in the relevant respects, equivalent. One option open to such explanations is to adopt an alternative, non-scalar inference analysis for the inferences children readily derive.

In fact, such a strategy is entertained in the study by Papafragou and Musolino (2003) to explain the high rate at which children computed an exactly interpretation of numerals. However, adopting such a strategy means abandoning the gains in parsimony achieved by explaining the derivation of so many, seemingly disparate, inferences through the one scalar inference process. Therefore, before resorting to such a position, it seems prudent to first seek an explanation that might account for children's variable behaviour in a way that retains a scalar inference analysis for the inferences in question. The second group of theories, called the *Alternatives-based approach*, seem to provide such explanations.

The Alternatives-based approach

The Alternatives-based approach attributes children's variable behaviour with scalar inferences to limitations that affect children's ability to generate the alternative sentences involved (Chierchia et al., 2001; Gualmini et al., 2001; Reinhart, 2006; Barner and Bachrach, 2010; Barner et al., 2011; Tieu et al., 2016; Singh et al., 2016). The Alternatives-based approach can be broken down into at least two distinct proposals, based on the specific limitation attributed to children.

One variant of the Alternatives-based approach proposes that children's behaviour is a result of limitations in their processing capacity (Chierchia et al., 2001; Gualmini et al., 2001; Reinhart, 2006; Tieu et al., 2016). This proposal contends that children do not possess the processing resources (e.g., working memory capacity) necessary to carry out the generation of alternative sentences, or to compare alternative sentences to the asserted sentence.

Another strand of the Alternatives-based approach contends that children's behaviour is the result of limitations in their knowledge of the abstract lexical scales involved in alternative sentence generation (e.g., <some...all>) (Barner and Bachrach, 2010; Barner et al., 2011). This proposal suggests that children do not have the lexical scale knowledge required to generate alternative sentences.

Both of these proposed limitations are predicted to result in children experiencing difficulties deriving scalar inferences. In this way, the Alternatives-based approach is able to match the Pragmatics-based approach in accounting for the low rate at which children derive many scalar inferences. However, the Alternatives-based approach goes even further, as it is also able to capture children's success in deriving certain scalar inferences (e.g., free choice inferences). More specifically, children have been found to succeed in deriving scalar inferences in cases where the linguistic or environmental context assists children in overcoming the proposed limitations. For example, we already mentioned that sentences like (16) are associated with free choice inferences like (17). However, we did not mention that free choice inferences like (17) are derived from alternative sentences that can be formulated from linguistic material contained in the asserted sentence in (16). Specifically, the free choice inference in (17) is derived from the alternative sentences in (18) and (19), which can both be formulated from linguistic material contained in the assertion in (16).

(16) Kung Fu Panda may push the green car or the orange car.

(17) Kung Fu Panda may push the green car and Kung Fu Panda may push the orange car.

(18) Kung Fu Panda may push the green car.

(19) Kung Fu Panda may push the orange car.

Similarly, ad-hoc inferences, like those Stiller et al. (2015) found children deriving readily, present the alternatives explicitly in the environmental context. Linguistic and environmental contexts with these properties plausibly reduce the processing and scale knowledge requirements associated with scalar inference derivation, because the alternative sentences can be composed from material presented explicitly in these contexts, rather than language users needing to retrieve this material from their mental lexicons.

Considering the scalar inferences that children have been found to derive readily, many of them present such linguistic material in their linguistic or environmental contexts (e.g., free choice inferences (Tieu et al. 2016, Chapter 3), ad-hoc inferences (Stiller et al., 2015), ignorance inferences (Hochstein et al., 2016), and distributive inferences (Chapter 2)). In this way, the Alternatives-based approach is able to account for many of the results where children have been found to derive scalar inferences readily.

In sum, the Alternatives-based approach is able to explain both children's typical difficulties with scalar inference computation, as well as many of the cases in which they succeed. Moreover, the Alternatives-based approach makes the following prediction regarding children's behaviour with certain scalar inferences that have not yet been investigated. If the linguistic material from which a scalar inference's alternatives are composed is presented within the asserted sentence, then children will derive that scalar inference successfully. In this chapter, we test this prediction by investigating children's behaviour with two such scalar inferences.

4.1.3 Testing the Alternatives-based approach

To test the Alternatives-based approach, we will investigate children's behaviour with two inferences associated with EverySome sentences like (20) - namely, the NotEvery inference in (21) and the None inference in (22).

(20) Every pig carried some of his rocks.

(21) Not every pig carried all of his rocks.

(22) None of the pigs carried all of his rocks.

We present more detail on how these inferences are derived as scalar inferences in the this chapter's appendix (Section 4.4.2 and Section 4.4.3). However, we will note here that, for both these inferences, the associated asserted sentence in (20) contains the linguistic material from which their alternative sentences are composed. For example, the NotEvery inference in (21) is derived as a scalar inference through the negation of the alternative sentence in (23). The alternative sentence in (23) is generated from the asserted sentence in (20) by replacing the existential quantifier 'some', with the universal quantifier 'all'. The universal quantifier 'every' is provided in the asserted sentence, thereby providing children with the linguistic material necessary to generate the alternative sentence in (23).

(23) Every pig carried all of his rocks.

As mentioned, when an asserted sentence possesses properties like these, the Alternatives-based approach predicts that children will be successful in deriving the relevant scalar inferences. Therefore, the Alternatives-based approach predicts that children will be successful in accessing inference-based interpretations of EverySome sentences.[6]

Note that there are important differences between EverySome sentences and the assertions associated with the inferences that children have previously been found to derive readily. The asserted sentences investigated in these previous studies present the relevant linguistic material as a more complete string (e.g., as one of the asserted sentence's individual disjuncts) (Hochstein et al. 2016; Tieu et al. 2016; Singh et al. 2016, Chapter 2, Chapter 3). In contrast, while the *EverySome* presents the critical linguistic material (i.e. the existential quantifier and the universal quantifier), a greater level of reconstruction is required to formulate the alternative sentences. To put it another way, it needs to be assumed that the content presented in the assertion can be 'recycled' and used in the generation of alternative sentences. For this reason, the current chapter extends this previous work, by investigating whether *just* presenting this key linguistic material is sufficient to help overcome the Alternative-based approach's proposed limitations. In other words, this investigation tests a 'stronger' version of the Alternatives-based approach.

4.1.4 Differences between adults' and children's interpretations

As just outlined, the Alternatives-based approach predicts that children will drive an inference-based interpretation of EverySome sentences. However, the Alternatives-based approach does not make any predictions regarding which specific

[6]As noted by an anonymous examiner, given that 'every' is a different lexical item from 'all', in order to claim that the relevant linguistic material is in some sense 'contained' in the asserted sentence we need to adopt some extra assumption regarding the nature of alternatives. For example, we could assume that the presentation of one universal quantifier (e.g., 'every') in an assertion allows children to access other universal quantifiers (e.g., 'all') when generating any associated alternative sentences.

Alternately, we could claim that the relevant alternative sentence that children (and perhaps adults) derive, when presented with an EverySome sentence like (20), is the sentence '*Every pig carried every of his rocks*'. In this case, all the relevant linguistic material would be contained in the asserted sentence in (20).

We do not need to prefer one these assumptions, however, the adoption of one of them (or some equivalent assumption) is required in order to claim our experiment as a test of the Alternatives-based approach.

inference this interpretation will be based on. There are other theories, however, which would appear to make predictions in this regard. One such proposal, presented in Crain et al. (1994), suggests that when interpreting sentences, adults and children are guided by different principles. In brief, children are proposed to be motivated by learnability considerations, which lead them to prefer stronger interpretations, while adults are motivated by processing considerations, which lead them to prefer weaker interpretations.

Crain et al. (1994) note that when children are presented with a sentence, they need to avoid the possibility of adopting an interpretation that is too weak. This is because, if they were to adopt such an interpretation, they would never be able to learn from positive evidence that an alternative, stronger interpretation was possible. For example, the two inferences we are investigating sit in a subset-superset relationship, where the conditions that would make the None inference true are a subset of the conditions that would make the NotEvery inference true. If children were to adopt the weaker, superset interpretation (i.e. the NotEvery inference in (25)) first, they would have no way of learning, from positive evidence, that the stronger, subset interpretation (i.e. the None inference in (26)) is also a possible interpretation of a sentence like (24). In order to avoid this situation, it is suggested that children first adopt the strongest available interpretation of a sentence. This idea is also discussed in a number of other papers and is often called the *semantic subset principle* (Berwick, 1985; Crain, 1992; Crain and Thornton, 1998; Moscati et al., 2016). Considering our target inferences, Crain et al. would predict that children should prefer the stronger None inference in (26) over the weaker NotEvery inference in (25).

(24) Every pig carried some of his rocks

(25) Not every pig carried all of his rocks.

(26) None of the pigs carried all of his rocks.

Crain et al. (1994) go on to propose that, in contrast, adults are motivated by different priorities. Specifically, that adults are motivated to prefer sentence representations that make the fewest commitments to their mental models. This is done to reduce cognitive effort, by lowering the chance that the interpretation will need to be revised. This idea has been called the *principle of parsimony* (Crain and Steedman, 1985). Considering our target inferences, Crain et al. would predict that adults should prefer interpretations including the weaker NotEvery inference in (25) over the stronger None inference in (26).[7]

A series of experiments have investigated adults' interpretations of EverySome sentences (Geurts and Pouscoulous, 2009a; Clifton Jr and Dube, 2010; Chemla and Spector, 2011; van Tiel, 2014). The results of these experiments, seem to suggest that while both the NotEvery and the None inferences are possible interpretations of EverySome sentences, the NotEvery inference interpretation seems to be preferred by adults. This finding is in line with Crain et al. (1994)'s expectations.[8] Children's interpretation of these sentences has, to our knowledge, not been investigated previously.

We should note that there is, of course, another even weaker possible interpretation of sentences like (24); namely, the literal interpretation that *every pig carried at least one of his rocks*. However, it is predicted that for many language users the derivation of an inference-based interpretation will be required by, for example, the general reasoning principles and conversational norms we outlined earlier (see Section 4.1). It is in these cases that Crain et al.'s proposals are predicted to apply.

In sum, Crain et al. (1994) propose that independent principles (i.e. the semantic subset principle and the principle of parsimony) govern adults' and children's default interpretations of sentences. For our target sentences these principles are predicted to lead adults and children to prefer different inference-based interpretations. That is, while children are predicted to prefer the None inference, adults are predicted to prefer the NotEvery inference.

[7]To avoid any confusion, we would like to note that while the None inference clearly entails the NotEvery inference, on some occasions of use, a language user could formulate the weaker NotEvery inference, without being committed to the stronger None inference. For example, consider a circumstance in which one of the pigs carried all of his rocks, but the others carried only some of their rocks. The sentence under consideration would be judged to be false by language users who generated the None inference, but it would be judged to be true by language users who generated the NotEvery inference. The experiment presented in this chapter distinguishes between these two inferences by presenting the test sentences in a series of different contexts, in order to pinpoint which of these inferences is computed by the child and adult participants.

[8]As we note and explore further in the next section (Section 4.1.5), this result has been interpreted by some papers as evidence that the NotEvery and None inferences are derived through different processes (Geurts and Pouscoulous, 2009a; Geurts and van Tiel, 2013; van Tiel, 2014).

Therefore, in addition to testing the predictions of the Alternatives-based approach, our investigation of these inferences will also provide an opportunity to test Crain et al. (1994)'s proposal regarding adults' and children's default interpretations.

4.1.5 Competing theories of scalar inference derivation

The Pragmatic and Grammatical accounts invoke different mechanisms to explain how scalar inferences are derived. According to the Pragmatic account, scalar inference derivation is based on general reasoning and conversational norms, whereas, the Grammatical account proposes that such inferences require the application of the silent grammatical operator, EXH. The mechanisms invoked by the Pragmatic account apply at the *whole-sentence* level, after the basic sentence meaning has been computed. By contrast, the EXH operator can be applied at both the whole-sentence level and sentence internally. For example, a sentence like (27) has two sites at which EXH can be applied, at the whole-sentence level, as in (28), and embedded under the universal quantifier, as in (29). The application of EXH at these different levels results in the derivation of different scalar inferences. The NotEvery inference in (30) is associated with the application of EXH at the whole-sentence level, and the None inference in (31) is associated with the application of EXH under the universal quantifier.

(27) Every pig carried some of his rocks.

(28) EXH[Every pig carried some of his rocks.]

(29) Every pig EXH[x carried some of his rocks.]

(30) Not every pig carried all of his rocks.

(31) None of the pigs carried all of his rocks.

In contrast, the mechanisms proposed by the Pragmatic account (i.e. general reasoning and conversational norms) are only applicable at the whole-sentence level, after the basic sentence meaning has been computed. Therefore, while the Pragmatic account is easily able to capture the derivation of the NotEvery inference in (30) as a scalar inference, capturing the None inference is not so straightforward.

One way for the Pragmatic account to capture the None inference is to suggest that the sentence in (32) should be included in the set of alternatives associated with an assertion like (27) (Spector, 2003; Sauerland, 2004; Chemla, 2009a). That is, the inclusion and subsequent negation of the sentence in (32) would result in the derivation of the inference in (33), which is equivalent to the None inference in (31).

(32) Some pig carried all of his rocks.

(33) It's not the case that some pig carried all of his rocks.

In order to admit the sentence in (32) to the relevant set of alternative sentences, however, the relevant explanations must include two additional assumptions in the scalar inference process. First, they must assume that alternative sentences can be composed by replacing multiple lexical items. This is necessary because the proposed alternative sentence in (32) is generated from the asserted sentence in (27) by replacing both the lexical items 'some' and the lexical item 'every'. That is, 'some' is replaced with 'all', and 'every' is replaced with 'some'. The second assumption such explanations must adopt is that alternative sentences are not required to asymmetrically entail the asserted sentence. This assumption is required because the alternative sentence in (32) is logically independent of the asserted sentence in (27). By adding both assumptions, certain proponents of the Pragmatic account are able to generate the None inference as a scalar inference (e.g., Spector 2003; Sauerland 2004; Chemla 2009a).[9]

Another way for the Pragmatic account to explain the None inference is to adopt a non-scalar inference analysis for it. This option is explored in Geurts and van Tiel (2013). They propose that the None inference and similar interpretations are derived by pragmatic *truth-conditional narrowing*. On this proposal, a hearer is, at times, led to infer that a speaker is adopting a non-conventional meaning of an asserted word or phrase. A hearer will often be guided to adopt such an interpretation by certain elements of the environmental or linguistic contexts. For example, Geurts et al. suggest that when presented with the sentence in (34), wherein the repetition of 'drinks' is marked with contrastive stress, a hearer would

[9]See Fox (2007) and Magri (2009) for discussion of issues related to the inclusion of these assumptions.

often be led to attach some additional meaning to the lexical item, beyond its conventional/dictionary meaning, e.g., *he drinks a lot*.

(34) When he drinks, he DRINKS.

Geurts et al. claim that, in a similar way, when a speaker produces an EverySome sentence like (27), a hearer may take it that they intend to convey the non-conventional None inference in (31). It is worth noting that Geurts et al. contend that interpretations based on truth-conditional narrowing are expected to be more difficult to access than those based on the scalar inference process (i.e. that they are 'marginal' interpretations). Moreover, they claim that assertions attempting to convey such interpretations are likely to be accompanied by some extra cue (e.g., contrastive prosodic stress) to signal that the assertion's meaning diverges from convention. As a result, the analysis presented by Geurts et al. expects that interpretations of EverySome sentences including the None inference should be less preferred than those including the NotEvery inference.

At this point we should remember that, as we mentioned in Section 4.1.4, a series of previous studies investigated and compared the rates at which adults derived the NotEvery inference and the None inference (see Section 4.1.4)(Geurts and Pouscoulous, 2009a; Clifton Jr and Dube, 2010; Chemla and Spector, 2011). These experiments found that the None inference was a possible, but less preferred, interpretation of EverySome sentences. That is, adults were found to access the NotEvery inference significantly more often than the None inference. These findings are consistent with the proposals by the Pragmatic account proponents - that the None inference is derived through a more complex instantiation of the scalar inference process (Spector, 2003; Sauerland, 2004; Chemla, 2009a), or that it is a marginal interpretation Geurts and van Tiel (2013).

However, as noted in Section 4.1.4, there is an alternative explanation of adults' preference for the NotEvery inference, based on the proposals of Crain et al. (1994). Crain et al. propose that a principle of parsimony entreats adults to prefer weaker interpretations, in order to avoid the cognitive effort associated with having to revise their interpretations. Therefore, it is not clear if the adults' preference for the NotEvery inference over the None inference is a result of the None inference being more difficult to access, as suggested by the Pragmatic account, or a result of it being the weaker inference-based interpretation, as suggested by Crain et al..

Investigating children's interpretations of EverySome sentences provides a unique opportunity to adjudicate between these two possible explanations of adults' preference for the NotEvery inference. This is because, according to Crain et al. (1994), children are not guided by the same parsing principles as adults. In fact, children are expected to initially exhibit preferences for stronger interpretations, which, in this case, would be the None inference. If, therefore, adults' documented preference for the NotEvery inference is a result of parsing principles, then, following Crain et al., we should not expected children's interpretations to be affected in the same way. However, if this preference is caused by the None inference being a more difficult to access interpretation, as suggested by the Pragmatic account, then children are expected to follow adults in preferring the NotEvery inference over the None inference. In fact, children might be expected to disfavour the None inference even more than adults do since children have been found to be less sensitive than adults to marked interpretations, including ones that are associated with contrastive stress (Choi and Mazuka, 2003; Snedeker and Yuan, 2008; Zhou et al., 2012).

Moreover, if children were found to prefer the None inference over the NotEvery inference, this would be consistent with the Grammatical account's proposal that these inferences are derived through the same process. That is, such a result is consistent with the suggestion that there is a high level of continuity between the nature of these inferences, and that any differences in language users' behaviour towards them is a result of interactions with other linguistic mechanisms (e.g., parsing principles), rather than any differences in how they are derived.

In sum, the Pragmatic and Grammatical accounts of scalar inferences are both able to capture the derivation of NotEvery and None inferences from EverySome sentences. The Grammatical account takes both inferences to be derived through essentially the same process (Chierchia, 2004, 2006; Chierchia et al., 2011; Chierchia, 2013; Fox, 2007). In contrast, the Pragmatic account takes them to be derived through somewhat different processes, and, as a result, expects the None inference to be a dispreferred (Spector, 2003; Sauerland, 2004; Chemla, 2009a; Geurts and van Tiel, 2013). The previous results showing that adults prefer interpretations including the NotEvery inference over the None inference are more consistent with proposals of the Pragmatic account - that the None inference is the more complex or marginal interpretation. However, it could be that adults' behaviour in these experiments was motivated instead by a parsing principle of parsimony favouring weaker inferences (Crain et al., 1994). Children, on the other hand, are expected to be unaffected by the principle of parsimony, and in fact, are proposed to be guided by the learnability considerations to prefer stronger interpretations, which in this case would be the None inference. Therefore, exploring children's interpretation of EverySome interpretations provides a unique opportunity to adjudicate between these accounts. The Pragmatic account expects for children to prefer the NotEvery inference, while, the Grammatical account (adopting Crain et al.'s assumptions) predicts that children will derive the None inference. We now turn to the experiment.

4.2 Experiment

The aim of our experiment was to investigate children's and adults' derivation of the NotEvery inference and the None inference, associated with EverySome sentences. Investigating these sentences provide us with the opportunity to explore the interesting possibility, raised by the Alternatives-based approach, that children will derive inference-based interpretations of EverySome sentences at an adult-like rate. Moreover, it allows us to test whether Crain et al. (1994)'s proposal extends to scalar inferences by investigating whether children will prefer None inferences over NotEvery inferences, while adults prefer the opposite. Finally, it provides the opportunity to adjudicate between the Grammatical and Pragmatic accounts of scalar inference derivation, by exploring whether, as the Pragmatic account contends, the None inference is a dispreferred interpretation of EverySome sentences. We investigated these theoretical issues through an experiment designed to test the following hypotheses:

- **Hypothesis 1:** Children will derive an inference-based interpretation of the target sentences as readily as adults.
- **Hypothesis 2:** Children's inference-based interpretations will be a result of having derived the None inference, whereas adults' inference-based interpre-

tations will be a result of having derived the NotEvery inference.

4.2.1 Method

Participants

Eighteen monolingual English-speaking adults (Macquarie University undergraduate students, 14 females, 4 males) and 31 monolingual English-speaking children (4;0-5;10, M = 4;05, 20 female, 11 males) were included in our experiment.[10] The adults took part in the experiment for course credit, or for a payment of $15.00. Children were recruited from several on-campus daycares at Macquarie University, and from a Macquarie University child research participant database. Informed consent was obtained from the adult participants, and from the parent/guardian of the child participants.

Procedures

This study used a *Truth Value Judgment Task* (Crain and Thornton, 1998). This task involves two experimenters. One experimenter presents a story, and the other wields a puppet who watches the stories along with the participant. Once a story is finished, the first experimenter asks the puppet to describe some aspect of the story they have just been told. The puppet responds with a target sentence.[11] The first experimenter then asks the participant whether what the puppet said was right or wrong. The participant responds with a *yes* or *no* judgment. If the participant provides a *no* judgment, then the experimenter asks the participant to provide a justification (i.e. "Why do you think <Puppet's name> is wrong?"/"What really happened?").[12] For adults, the materials were the same, however, they were asked to write down their judgments and justifications, and to include justifications for both *yes* and *no* judgments.

[10] As we outline in Section 4.2.1, this experiment is comprised of two sessions. These ages correspond to the age of the child participants at the first of these two sessions.

[11] Note that for adult participants the puppet's sentences were pre-recorded and presented to participants at the appropriate times via a laptop. Child participants were given the traditional live presentation of the sentences via a second experimenter.

[12] Justifications were not requested for *yes* responses, as such a request is infelicitous and so may confuse child participants (Crain and Thornton, 1998).

Children were tested individually, either in the lab or in a quiet room at their daycare. Adults were tested in groups of 1 to 3. The items were split across two sessions, and the sessions were conducted 7-14 days apart. Each session lasted approximately 20 minutes.

Materials

The experiment included four test conditions, two control conditions, and one filler condition.

Test conditions

Each item in our four test conditions consisted of a story involving three characters (e.g., pigs). Each of these three characters had a set of four objects (e.g., rocks) placed on an orange square in front of them. Each character could decide how many (if any) of their objects they would act upon (e.g., carry). The experimenter took on the role of each of the characters in turn. Each character went through a process of, considering their objects, deciding how many they wanted to act upon, and then doing so. Once all the characters had gone through this process, the experimenter asked the puppet what had happened in the story, to which the puppet responded with the test sentence. To illustrate, one of the items from the OneInference condition (described in more detail below) is presented in (35). This story would be associated with the test sentence in (36). The final scene of this story is shown in Figure 4.1.

(35) Example OneInference condition item.
This is a story about three pigs. These pigs each have rocks that they can carry if they want to. Let's see what they do:
Pig 1: *"Let me see, I'll carry this rock [carries rock (1/4)], and this one [carries rock (2/4)], should I stop? Hmm...I'm feeling really strong today, so I'll also carry this rock [carries rock (3/4)], and this rock too [carries rock (4/4)]."*
Pig 2: *"Let me see, I'll carry this rock [carries rock (1/4)], and this one [carries rock (2/4)]. should I stop? Yes I will, as I am tired.*

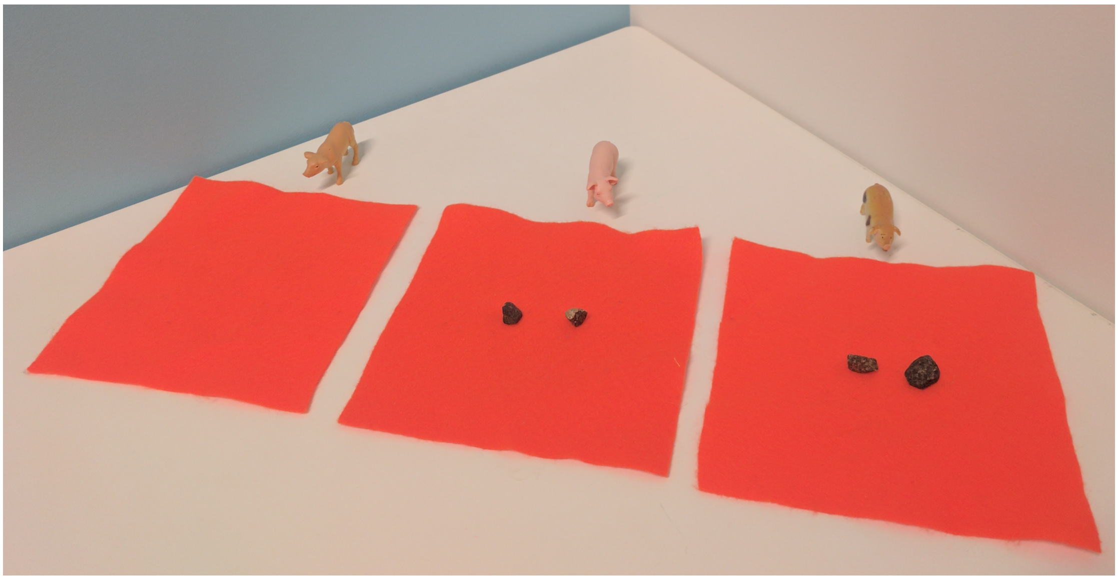

Figure 4.1: Final scene of the story presented in (35). The rocks remaining on the squares are those that have not been carried.

Pig 3: *"Let me see, I'll carry this rock [carries rock (1/4)], and this one too [carries rock (2/4)], should I stop?. Yes I will, as I am tired. [see Figure 4.1]*

Experimenter: Ok <Puppet's name>, what happened in that story?

(36) **Puppet:** Hmm, every pig carried some of his rocks.

There were three interpretations of our test sentences that we were interested in. First, there was the Literal interpretation in (38). Next, there was the NotEvery interpretation in (39), comprised of the Literal interpretation and the NotEvery inference. Finally, there was the None interpretation in (40), comprised of the Literal interpretation and the None inference.

(37) Every pig carried some of his rocks.

(38) **Literal interpretation:** Every pig carried at least one of his rocks.

(39) **NotEvery interpretation:** Every pig carried at least one of his rocks AND Not every pig carried all of his rocks.

(40) **None interpretation:** Every pig carried at least one of his rocks AND None of the pigs carried all of his rocks.

Our test conditions had a certain entailment relationship to these possible interpretations. That is, as shown in Table 4.1, in the False condition, none of the characters acted on any of their objects. Therefore, the False condition was inconsistent with all possible interpretations in (38)-(40). In contrast, in the True condition, all of the character acted on some but not all of their objects. Therefore, the True condition was consistent with all possible interpretations in (38)-(40). In the NoInference condition every character acted on all of their objects, making this condition consistent with the Literal interpretation in (38), but inconsistent with both the NotEvery interpretation in (39) and the None interpretation in (40). Finally, in the OneInference condition, two characters acted on two of their four objects, and one character acted on all of their objects. Therefore, the OneInference condition was consistent with the Literal interpretation in (38) and the NotEvery interpretation in (39), but was inconsistent with the None interpretation in (40).

Test Conditions	**Context**	**Interpretations**
True	2/4, 2/4, 2/4	Literal, NotEvery, None
OneInference	2/4, 2/4, 4/4	Literal, NotEvery, ~~None~~
NoInference	4/4, 4/4, 4/4	Literal, ~~NotEvery~~, ~~None~~
False	0/4, 0/4, 0/4	~~Literal~~, ~~NotEvery~~, ~~None~~

Table 4.1: Relationship between test conditions and relevant interpretations.

By comparing participants' responses between our four test conditions, we could determine the different interpretations participants were accessing. A difference in target sentence rejection rates between the False and the NoInference conditions is evidence that participants are accessing a Literal interpretation. A difference in target sentence rejection rates between the NoInference and OneInference conditions is evidence that participants are accessing a NotEvery interpretation. Finally, a difference in target sentence rejection rates between the OneInference and True conditions is evidence that participants are accessing a None interpretation.

Control and filler conditions

Our experiment also included two control conditions. These control conditions were designed to ensure that participants understood the basic meaning of the universal quantifier 'every'. That is, the Every_True and Every_False control conditions outlined in Table 4.2 were described with a control sentence like (41).

(41) **Sentence:** Every pig carried rocks.

Control Conditions	Context
Every_True	2/4, 2/4, 4/4
Every_False	2/4, 2/4, 0/4

Table 4.2: Control condition properties.

In addition to these test and control conditions, we also included a Filler condition. The items in the Filler condition were designed so that they had two possible target sentences for each item. One of which was designed to elicit a *no* response, the other, a *yes* response. The sentence that was ultimately used was chosen based on a participant's responses to previous trials, so that we could reduce the possibility of participants providing more than two *yes* or *no* responses in a row. For example, if a participant had rejected the target sentences of the two items preceding a Filler item, then in the Filler item, the filler sentence designed to elicit a *yes* response would be used.

Design

Participants were presented with all of these items over the course of two sessions. The conditions were split up between these sessions, in the manner outlined in (42).

(42) a. **Session A:** True, NoInference, Every_True, Every_False, Filler

b. **Session B:** OneInference, False, Filler

The ordering of the items within each session was pseudo-randomised by first creating a random order, and then slightly modifying it to ensure that, if participants responded as expected in the True and False test conditions, the Every_True and Every_False control conditions, and the Filler conditions, they would not accept or reject more than two target sentences in a row. A second version of each session was created, with the order of the trials reversed. The version of the sessions as well as the order in-which the sessions were presented was counter-balanced across participants.

4.2.2 Results

In order to be included in the final data-set participants needed to meet two exclusion criteria: First, participants needed to answer correctly at least 4 of the 5 items in the Filler condition. This was to ensure that participants understood the task, and were paying a minimal amount of attention to the stories and target sentences. All of the adult and child participants met this criterion. Second, participants were required to answer correctly at least 3 of the 4 items in Every_True and Every_False control conditions. This was to ensure that participants understood the basic meaning of the universal quantifier 'every'. All of the adult participants met this criterion, however, 10 of the children did not, and so were excluded from the final data-set. The final data-set, then, was comprised of 18 adult and 21 child participants.

Participant responses came in the form of judgments and justifications. We will consider each of these in turn.

Judgments

The judgment results are presented in Figure 4.2. To test for differences between conditions for each of our groups we ran a series of Wilcoxon signed-rank tests. To test for differences between groups in each of our conditions we ran a series of Wilcoxon rank-sum tests. Once we had generated the p-values for these tests, we used the Holm-Bonferroni procedure (Holm, 1979) to determine which contrasts were significant at an alpha value of 0.05. The results of the Wilcoxon signed-rank tests are presented in Table 4.3. The results of the Wilcoxon rank-sum tests are presented in Table 4.4.

Considering Table 4.3 and Figure 4.2 together, we found that children were

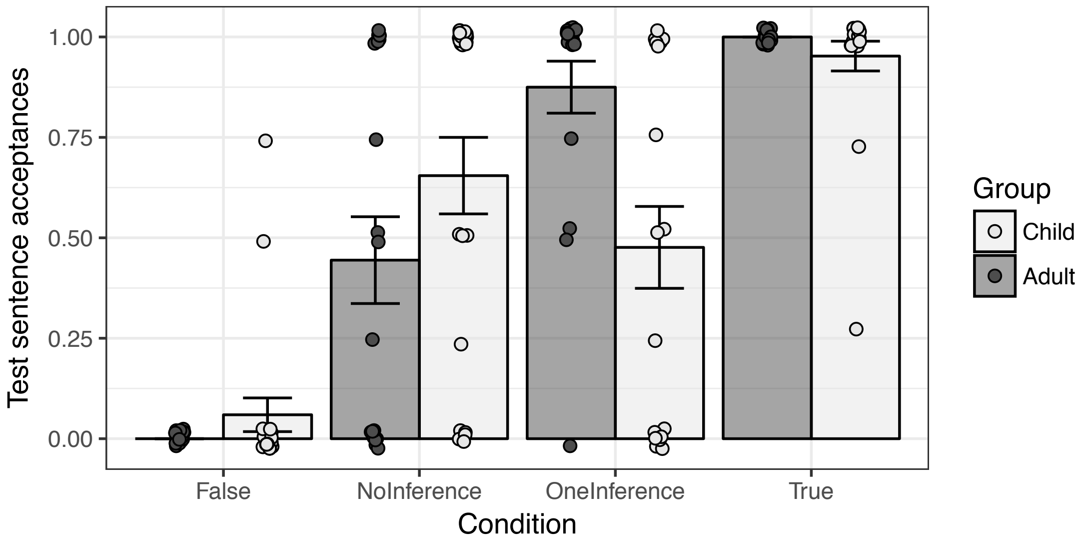

Figure 4.2: Mean proportion of test sentence acceptances across our test conditions. The vertical lines represent the standard error. Each dot represents one participant's mean test sentence acceptance rate for that condition. A horizontal jitter of .1 and vertical jitter of .025 were applied for easier visualisation.

accepting target sentences significantly more in the False condition, compared to the NoInference condition. We also found that children were accepting target sentences significantly more in the OneInference condition, compared to the True condition. However, we found no significant difference between children's acceptance of target sentences in the NoInference condition and the OneInference condition. As for adults, we found they were accepting target sentences significantly more in the False condition, compared to the NoInference condition. We also found that adults were accepting target sentences significantly more in the OneInference condition, compared to the NoInference condition. However, we found no significant difference between adults' acceptance of target sentences in the OneInference condition and the True condition.

The results of the Wilcoxon rank-sum tests, to investigate differences between groups in the different conditions are presented in Table 4.4.

Considering Table 4.4 and Figure 4.2, we found no significant difference between children and adults in the False condition, the NoInference condition, or the True condition. However, we did find that adults were accepting target sentences

Group	Comparison	Test statistic	Effect size	p-value
Child	FALSE vs. NOINF.	Z = -3.6	r = -.56	p = .003*
	NOINF. vs. ONEINF.	Z = -1.98	r = -.31	p = .048
	ONEINF. vs. TRUE	Z = -3.21	r = -.5	p = .001*
Adult	FALSE vs. NOINF.	Z = -2.82	r = -.47	p = .005*
	NOINF. vs. ONEINF.	Z = -2.92	r = -.49	p = .004*
	ONEINF. vs. TRUE	Z = -1.66	r = -.28	p = .098

Table 4.3: Results of Wilcoxon signed-rank tests. *Significant at α = .05, based on the Holm-Bonferroni correction procedure.

Condition	Comparison	Test statistic	Effect size	p-value
FALSE	Child vs. Adult	W = 171	r = -.21	p = .197
NOINF.		W = 139.5	r = -.24	p = .137
ONEINF.		W = 277	r = -.44	p = .006*
TRUE		W = 207	r = -.21	p = .197

Table 4.4: Results of Wilcoxon rank-sum tests. *Significant at α = .05, based on the Holm-Bonferroni correction procedure.

significantly more than children in the ONEINFERENCE condition.

These results lead us to the following conclusions regarding how participants were engaging with the target interpretations in (43)-(45).

(43) **Literal interpretation:** Every pig carried at least one of his rocks.

(44) **NotEvery interpretation:** Every pig carried at least one of his rocks AND not every pig carried all of his rocks.

(45) **None interpretation:** Every pig carried at least one of his rocks AND no pig carried all of his rocks.

The difference we found in target sentence acceptances for both adults and children, between the FALSE and NOINFERENCE conditions, is evidence of both groups accessing the Literal interpretation in (43). The difference we found in target sentence acceptances for adults between the NOINFERENCE and ONEINFERENCE conditions, is evidence of adults accessing the NotEvery interpretation in (44). In contrast, the lack of a difference in target sentence acceptances for children between the NOINFERENCE and ONEINFERENCE conditions, means that we do not have any evidence of children accessing the NotEvery interpretation in (44). The difference in target sentence acceptances for children between the OneInference and the True conditions is evidence of them accessing the None interpretation in (45). The absence of a significant difference in adult acceptances of the target sentences between the OneInference and True conditions means that we lack evidence that adults accessed the None interpretation in (45).

These results reveal both similarities and differences between our participant groups. Both groups accessed Literal and inference-based interpretations (i.e. the NotEvery interpretation or the None interpretation) of EverySome sentences at similar rates. However, the different groups accessed different inference-based interpretations. While children accessed the None interpretation, adults accessed the NotEvery interpretation. This conclusion is corroborated by the different acceptance rates between participant groups in the OneInference condition, as this was the only condition where these different inference-based interpretations would have resulted in different judgments.

To further explore the response patterns of individual participants, we attempted to identify which of the target interpretations each participant had accessed. To do this, we first identified the type of response each participant tended to produce in the NoInference and OneInference conditions. For example, if a participant accepted at least 3 of the 4 items in the NoInference condition, then they were categorised as 'accepting' the test sentences in that condition. Following this, we identified how many participants displayed the combination of responses expected for each target interpretation, as laid out in Table 4.5. For example, if a participant was categorised as accepting test sentences in both the NoInference and the OneInference conditions, they were categorised as having accessed a Literal interpretation. Twelve participants (8 children and 4 adults) provided mixed responses in at least one of these test conditions, and so they were not included in any of the interpretation categories.

For the 27 participants (13 children and 14 adults) who did provide consistent responses, Figure 4.3 presents the results of this categorisation. As Figure 4.3 shows, both groups were similar in the rates at which they accessed a Literal interpretation of our target sentences. However, they differed in regards to the inference-based interpretation they accessed, with children only accessing a None

	NoInference	OneInference
Literal	*accept*	*accept*
NotEvery	*reject*	*accept*
None	*reject*	*reject*

Table 4.5: Responses expected by target interpretations.

interpretation, and adults tending to access a NotEvery interpretation.

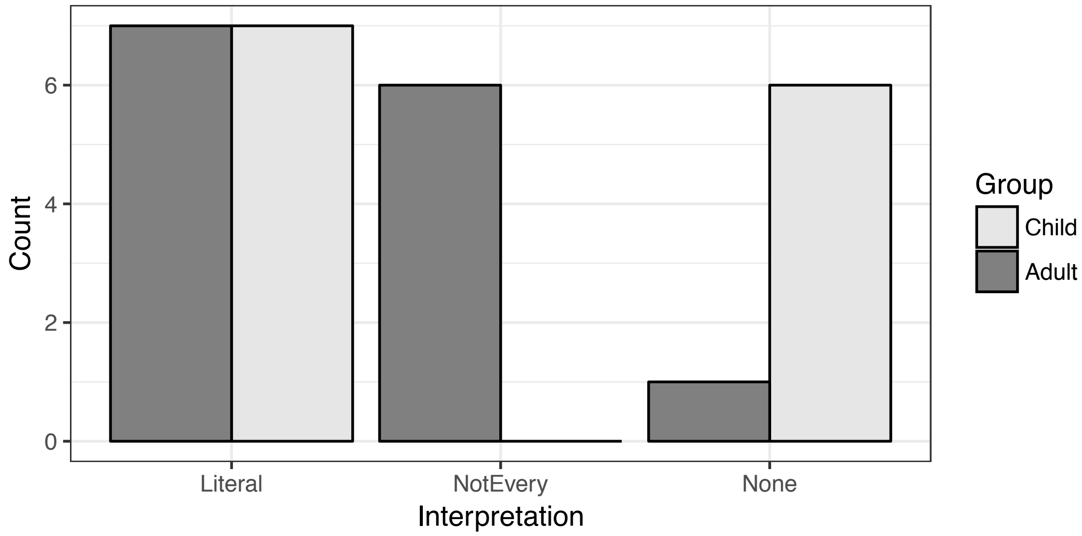

Figure 4.3: Participants categorised into different interpretations.

Therefore, this investigation into the response patterns of individual participants provides further corroborating evidence for the suggestion that children's inference-based interpretations were a result of having accessed a None interpretation, while adults' were a result of having accessed a NotEvery interpretation.

Justifications

We also elicited and recorded justifications for many of our participants' judgments. A number of considerations affected our analysis of this data. First, we focused only on justifications given in the NoInference and OneInference conditions, as these were the conditions that distinguished between the target interpretations. Second, we only included participants' justifications for target sentence rejections as we did not systematically elicit justifications from children when they accepted a target sentence. This meant that we were unable to obtain justifications that were associated with the Literal interpretation indicated in (46) below. Third, for each group, we only considered the condition that was inconsistent with the inference-based interpretation they appeared to have derived. This was because we were primarily interested in whether each group's justifications would be consistent with the inference-based interpretation attributed to them. In this regard, our group- and individual-level analyses suggested that children were primarily deriving the None interpretation in (48), while adults were primarily deriving the NotEvery interpretation in (47). Taking all these considerations into account, we analysed children's justifications for test sentence rejections in the OneInference condition, as well as adults' justifications for test sentence rejections in the NoInference condition.

(46) **Literal interpretation:** Every pig carried at least one of his rocks.

(47) **NotEvery interpretation:** Every pig carried at least one of his rocks AND not every pig carried all of his rocks.

(48) **None interpretation:** Every pig carried at least one of his rocks AND none of the pigs carried all of his rocks.

First we will consider the justifications children provided for their rejections of test sentences in the Oneinference condition. There were 44 justifications of this kind. These justifications were coded into one of three categories, shown in Figure 4.4. The bulk children's justifications (36/44; 82%) were coded as falling into the 'OneAll' category. The justifications in this category focused on the fact that one of the characters had acted on all of their objects (e.g., *No, because this dog ate all of them; No, that sheep carried all of the rocks; No, because this one used all of them*). These justifications focused on the one character that had acted on all of their items. These justifications support the suggestion that these participants were accessing the None interpretation in (48). That is, these participants were focusing on the aspect of the context that was not consistent with this interpretation.

The next more frequent type of justification (7/44; 16%) was the 'TwoNotAll' category. The justifications that we assigned to this category focused on the fact that two of the characters had not acted on all of their objects (e.g., *No, because these two didn't carry all of their rocks, like this one; No, these two didn't; No, because these two didn't use all of them*). These justifications seem to indicate that the participants assigned some alternative type of interpretation. Specifically, these justifications suggest that these child participants expected all of the characters to act on all of their items. Finally, the smallest number of justifications (1/44; 2%) were assigned to the category 'Other'. This justification merely repeated what had happened in the story, and so did not provide any insight into the specific interpretation this participant was accessing.

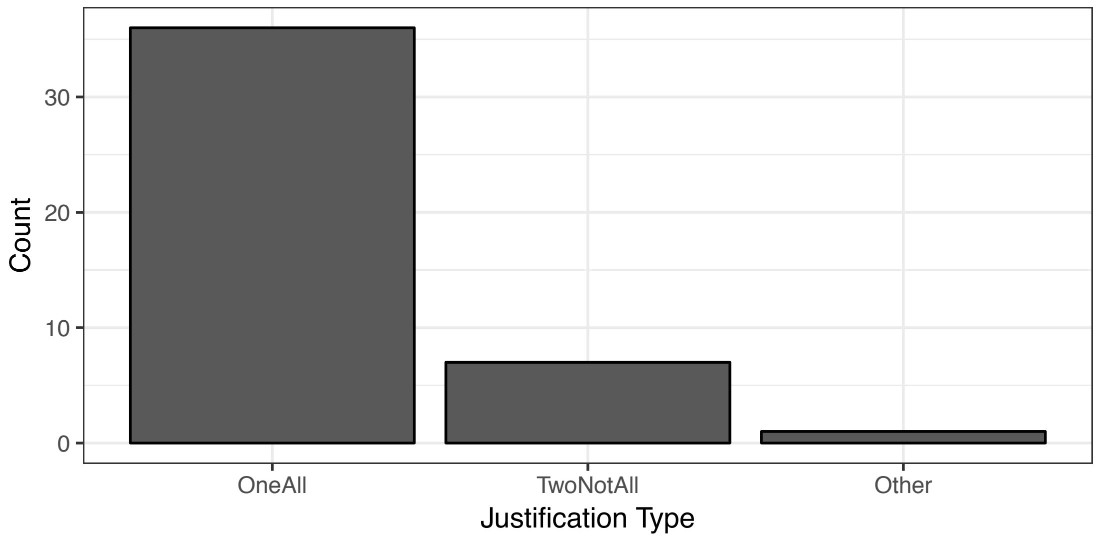

Figure 4.4: Justifications children provided for rejection test sentences in the One-Inference condition.

Now we will look at the justifications adults provided for their rejections of the test sentences in the NoInference condition. There were 40 justifications of this kind. These justifications fell into two categories, shown in Figure 4.5. The vast majority of justifications (36/40; 90%) were coded into the 'AllAll' category. These justifications focused on the fact that all of the characters had acted on all of their objects (e.g., *No, each of the unicorns polished all of their gemstones; No, every cat threw all of his glowsticks; No, the rabbits used all of their tea-bags to make their pots of tea*). The focus these AllAll justifications put on the fact that all of the characters acted on all of their objects provides support for the conclusion that these participants' were accessing a NotEvery interpretation of the test sentences. That is, these justifications focused on the aspect of the context that was inconsistent with the NotEvery interpretation in (47). The remaining justifications were coded into the category 'Other' category (4/40; 10%). These justifications merely repeated the basic elements of the story, for example mentioning that all of the characters had acted on their objects, without explicitly identifying that that had acted on 'all' of them (e.g., *No, all the cats threw glowsticks; No, every rabbit used teabags to make tea; No, each unicorn polished her stones*). These justifications did not provide any clear insights into the interpretation participants were accessing.

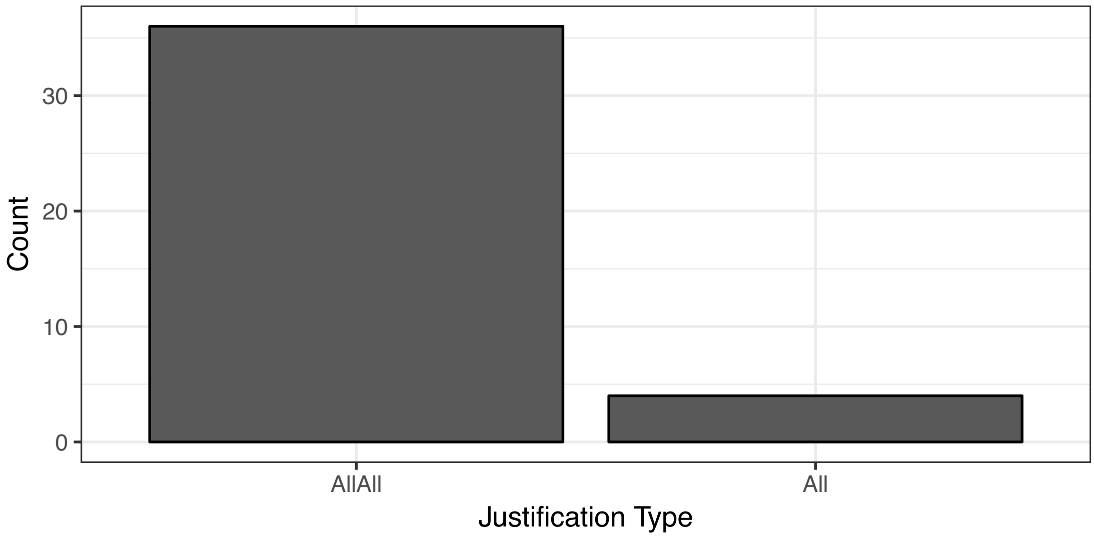

Figure 4.5: Justifications children provided for rejection test sentences in the One-Inference condition.

In sum, a coding analysis of the relevant justifications for each participant group underscores the conclusion that adults tended to access the NotEvery interpretation of test sentences, whereas children tended to access the None interpretation.

4.3 Discussion

The primary aim of our experiment was to further test the Alternatives-based approach. The Alternatives-based approach attributes children's variable success in scalar inference computation to limitations in their ability to generate alternative sentences. Consequently, the Alternatives-based approach predicts that children will more easily compute scalar inferences when the linguistic material from which the associated alternative sentences are composed is presented as part of the asserted sentence.

In relation to the present experiment, the Alternatives-based approach predicts that children will successfully access one of the inference-based interpretations associated with EverySome sentences. This is because the ingredients for constructing the relevant alternative sentences, the existential quantifier 'some' and the universal quantifier 'every' are contained in the asserted sentence.

We tested the Alternative-based approach's prediction by investigated the rates at which children and adults derived NotEvery inferences like (50) and None inferences like (51) from EverySome sentences like (49). Specifically, we wanted to test whether children would derive an inference-based interpretation of EverySome sentences at an adult-like rate.

(49) Every pig carried some of his rocks.

(50) Not every pig carries all of his rocks.

(51) No pig carried all of his rocks.

Our experiment found that children were deriving an inference-based interpretation of our target EverySome sentences at the same rate as adults, consistent with the Alternatives-based approach's expectations.

This result is particularly interesting when considered in light of the previous highly-replicated finding that, when presented with sentences like (52), children tend not to derived the associated OnlySome scalar inference in (53). Especially given that sentences like (52) are, in a sense, simpler versions of our test sentences (Noveck 2001; Papafragou and Musolino 2003; Guasti et al. 2005; Bill et al. 2016, among others).

(52) The pig carried some of his rocks.

(53) The pig didn't carry all of his rocks.

The contrast between these previous findings and our results provides even more support for the Alternatives-based approach's proposal that children's documented variation in scalar inference computation is due to limitations in how they interact with alternative sentences. To put it another way, when the Alternatives-based approach's proposed limitations are overcome by presenting certain content in the linguistic context, then, even if the asserted sentences and associated inferences are more 'complex', children will successfully compute the inferences anyway.

In sum, the findings of our experiment are consistent with the Alternatives-based approach's prediction that children would access an inference-based interpretation of EverySome sentences at the same rate as adults. This supports the Alternatives-based approach as a viable account of children's behaviour with scalar inferences.

4.3.1 Principles of interpretation

A secondary aim of our study was to investigate the proposal presented in Crain et al. (1994) that when interpreting sentences adults and children are guided by different principles. Specifically, Crain et al. propose that, while adults are guided by a principle of parsimony (Crain and Steedman, 1985), which leads them to prefer weaker interpretations, children are guided by a semantic subset principle (Berwick, 1985), which leads them to prefer stronger interpretations. These principles are not expected to affect whether or not language users access a literal or inference-based interpretation. However, it is plausible that, when a language user 'decides' to access an inference-based interpretation, these principles would influence which inference this interpretation is based on. That is, when deriving an inference-based interpretation, these principles should lead children to prefer the stronger None inference, while adults are predicted to prefer the weaker NotEvery inference.

Turning to our results, we found that children's inference-based interpretations tended to be a result of having derived the None inference, whereas the adult's inference-based interpretations tended to be a result of having derived the NotEvery inference. These findings are consistent with the proposal outlined in Crain et al. (1994).

Again, the reason why children are proposed to prefer stronger initial interpretation is so that they can be easily be falsified if that interpretation is not consistent with the interpretation assigned by adults. In contrast to children, adults are expected to generate weaker initial interpretations due to parsing considerations that motivate them to reduce the chance they will need to bear the cognitive processing cost associated with having to revise their interpretation.

Given the striking nature of these results, however, it seems prudent to first consider whether there might be some other explanation for our participants', and particularly children's, behaviour. For example, the following argument could be advanced to suggest that, instead of the subset principle, children's behaviour was motivated by a desire for the characters in our stories to behave uniformly.

The OneInference condition was the only condition where the characters' behaviour was not uniform. That is, in the OneInference condition, while two of the characters act on 2 out of their 4 objects, one character acts on 4 of their 4 objects. In contrast, in every other test condition (i.e. False, NoInference, True) each character acts upon the same number of objects. Therefore, if, as a result of the test sentences including a universal quantifier, children had wanted the characters' actions to be uniform, then this would also have generated our finding of children rejecting the target sentences in the OneInference condition. It could be argued then, that children's behaviour was motivated by a desire for characters actions to be uniform, rather than the subset principle appealed to by Crain et al. (1994).

We take such an explanation of children's behaviour in the OneInference condition to be implausible for two reasons. First, as already mentioned, in contrast to the OneInference condition, the NoInference condition presents stories wherein the characters' behaviour was uniform, with each character acting upon the same number of objects (4 out of 4). Despite, this, children rejected test sentences in the NoInference condition at a similar rate to the OneInference condition. If uniformity considerations were the only reason children rejected the OneInference condition, children should not have also rejected the target sentences in the NoInference condition.

Second, our exclusion criteria required participants to correctly respond to at least 3 out of the 4 Every control items. The two Every_True control condition items presented exactly the same context as the OneInference condition, but were described with a target sentence like (54), which included a universal quantifier in the same position as our test sentences. Therefore, the participants included in our analysis had accepted at least one target sentence like (54), as a description of a non-uniform context. If children required characters to be uniform in their behaviour, we would not expect them to accept the target sentences in this condition.

(54) Every pig carried rocks.

Another possible explanation for our results (as noted by an anonymous examiner) arises from the fact that children's rejection of test sentences in both the NoInference and the OneInference conditions was around 50%. That is, one could propose that children in our experiment were confused by those conditions, and so were merely rejecting the test sentences randomly. However, if we consider the justifications children provided, the majority of rejections (82%) in the OneInference condition were associated with a clear and coherent justification. Moreover, this justification focused on the fact that one of the characters had acted on all of their items, precisely the kind of justification we would expect if they were accessing the None interpretation. For this reason, we believe it is implausible that children's responses in these conditions are a result of confusion. Instead, we take it that the bulk of these children were accessing a genuine None interpretation of our test sentences.

In sum, despite the striking nature of this result, the most plausible explanation of it seems to be the proposal presented by Crain et al. (1994), that adults' and children's sentence interpretations are guided by different principles. Namely, that children are guided by learnability considerations (i.e. the subset principle), whereas adults are guided by processing reduction considerations (i.e. the principle of parsimony).

4.3.2 Adjudicating between accounts of scalar inferences

Our experiment also provided a unique opportunity to adjudicate between two contemporary accounts of scalar inference derivation - the Grammatical account and the Pragmatic account. The Grammatical account captures the derivation of the NotEvery inference and the None inference through essentially the same process (i.e. the application of EXH to difference parts of the asserted sentence).

The Pragmatic account's relationship with these inferences is more complicated. While it can also take the NotEvery inference to be derived through the classical scalar inference process, it is not able to capture the None inference in the same way. This is because, unlike the Grammatical account, the Pragmatic account contends that the scalar inference process can not be applied sentence internally. Therefore, it is more difficult for the Pragmatic account to capture the None inference as a a scalar inference.

One way for the Pragmatic account to generate the None inference is by amending the scalar inference process to allow alternative sentences to be composed by replacing multiple lexical items in the asserted sentence, and by allowing alternative sentences to be logically independent from asserted sentences (Spector, 2003; Sauerland, 2004; Chemla, 2009a). While this explanation still appeals to the scalar inference process, the fact that the alternative sentence is generated through multiple lexical replacements means that the process is more complex than the NotEvery inference's derivation.

Another way for the Pragmatic account to capture the None inference is to propose an alternative mechanism altogether, such as the process of pragmatic truth conditional narrowing - proposed in Geurts and van Tiel (2013), which derives the None inference as a marginal interpretation. For present purposes, the important thing to note is that, regardless of the specifics, the Pragmatic account takes the NotEvery inference to be more difficult to access than the None inference. In contrast, the Grammatical account takes both of these inferences to be derived through essentially the same process, and so expects these inferences to be more equal in their accessibility.

Earlier we noted that the findings of previous experiments investigating how adults derive the NotEvery and None inferences appeared to favour the Pragmatic account. In these studies, adults were found to systematically prefer the NotEvery inference over the None inference (Geurts and Pouscoulous, 2009a; Clifton Jr and Dube, 2010; Chemla and Spector, 2011). However, an alternative explanation of this adult preference can be advanced, based on the proposal we just presented by Crain et al. (1994). These researchers proposed that adults preferences are based on a principle of parsimony, that favours weaker sentence interpretations. However, children are not expected to be influenced by the principle of parsimony, due to the learnability problems that would ensue. Instead, Crain et al. propose that children are guided by the subset principle, which leads them to initially adopt stronger sentence interpretations. Assuming this proposal then, investigating children's derivation of these inferences provides a unique opportunity to avoid some of the confounds that limit the conclusions that can be drawn from adults' behaviour. One the one hand, if the Grammatical account is right and both these inferences are derived through essentially the same process, then children's inference-based interpretations are expected to be from having derived the None inference. On the other hand, if the Pragmatic account is right and the None inference is derived through a more complex process, or is a marginal interpretation, then children are expected to follow adults in preferring the NotEvery inference.

The present experiment found that children's inference-based interpretations were based largely on the None inference. Based on the reasoning just presented, this result is consistent with the expectations of the Grammatical account, and so provides further support for it as an explanation of how scalar inferences are derived.

4.3.3 Children removed by exclusion criteria

As mentioned earlier (see Section 4.2.2), 10 children were excluded from our final dataset because they failed to pass at least 3 out of the 4 control condition items. Given the size of this group, we looked more closely at these participants responses and justifications to see if they revealed any insights into what might have been motivating their behaviour. Two of these children produced responses suggesting they did not have a target understanding of the universal quantifier's contribution to the test sentence (e.g., they wanted every character to have acted on all of their objects). Three more of these children were unwilling to reject any of the puppet's target sentences. However, the remaining 5 children produced a more interesting pattern of responses. These 5 children were excluded as a result of rejecting both of the Every_True control items. Their justifications for these rejections tended to describe in greater detail what had happened in the story (e.g., *(No) One ate all of them, and two of them ate two leaves; (No) That one only ate two, and that one only ate two, and that one ate all of them; (No) Two donkeys ate some of the* *leaves and one donkey ate all of them*). Moreover, when we looked at the other conditions, these participants also tended to reject the test sentences across all the test conditions, including the True condition, which was consistent with all of our target interpretations. The justifications these children provided for rejecting these test sentences were similar to those provided for Every-True condition rejections. That is, they also tended to focus on the fact that there was a more informative way to describe the relevant context - namely, by using numerals in place of the existential quantifier *some*. For example, these participants tended to justify their True condition rejections by focussing on the fact that the characters had all acted on *two* of their objects (e.g., *(No) Because two; (No) They each used two stars; (No) They all ate two of them*).

This pattern of judgments and justifications seem to suggest that this group of children rejected the target sentences because there was a more informative way to describe the context. That is, rather than describing the number of items characters had acted on using the existential quantifier 'some', these participants wanted the target sentences to be more specific by mentioning the specific number of objects that each character had acted upon. None of the adult participants adopted this response strategy. Therefore, it would appear that this subgroup of child participants had even more rigid requirements than adults did regarding the informativity of the target sentences. Notably, the ability to judge the informativity of an asserted sentence relative to alternative sentences is a necessary precursor to the ability to derive scalar inferences. Moreover, it is an ability that the Alternatives-based approach expects children to have (Chierchia et al., 2001; Papafragou and Musolino, 2003; Guasti et al., 2005).

The fact that children seem to be rejecting the target sentences on the basis of more informative sentences including numerals is also interesting given that previous research has found that children's knowledge of the numeral scale seems outstrip that of other scales (Papafragou and Musolino, 2003; Barner and Bachrach, 2010; Huang et al., 2013). It could be that these children take numerals and quantifiers to be competitors in a way that adults do not.

In sum, it would appear that half of the children that we removed from our data set through our exclusion criteria were nonetheless responding to our sentences in a consistent, although non-adult-like manner. That is, these child participants appeared to have higher informativity requirements than adults, wanting the target sentences to include a specific numeral in place of the more vague 'some' quantifier. This pattern is consistent with previous results suggesting that children are sensitive to the relative informativity of sentences (Chierchia et al., 2001; Gualmini et al., 2001; Papafragou and Musolino, 2003; Guasti et al., 2005), and that the numeral-scale is learnt early in development (Fuson, 1988; Papafragou and Musolino, 2003; Barner and Bachrach, 2010; Huang et al., 2013). Further exploration of this behaviour could be a promising area for future research.

4.3.4 Conclusion

The Alternatives-based approach represents a group of explanations that attribute children's variable success in computing scalar inferences to certain limitations affecting how they interact with the alternative sentences involved (Chierchia et al., 2001; Gualmini et al., 2001; Reinhart, 2006; Barner and Bachrach, 2010; Barner et al., 2011; Tieu et al., 2016). This approach predicts that children will more easily compute scalar inferences when linguistic material from which the relevant alternative sentences are formulated is contained in the asserted sentence. We tested this prediction by comparing the interpretations that children and adults accessed for EverySome sentences. Two scalar inferences have been associated with these sentences - the NotEvery inference and the None inference (Sauerland, 2004; Chierchia, 2004; Fox, 2007; Chemla, 2009a; Geurts and van Tiel, 2013). For both these inferences, the linguistic material from which the relevant alternative sentences are composed is contained in the asserted EverySome sentence. Therefore, the Alternatives-based approach predicts that children will be successful in accessing an inference-based interpretation of EverySome sentences. Our results were consistent with this prediction, providing further support for the Alternatives-based approach as an account of children's variable performance with scalar inference computation.

The Alternatives-based approach does not make a distinction between the NotEvery and the None inference as far as children's predicted behaviour is concerned, however, a proposal outlined in Crain et al. (1994) is not so agnostic. Crain et al. propose that when interpreting sentences adults and children are guided by different principles. Specifically, that children are guided by a semantic subset principle which leads them to prefer stronger interpretations. Whereas adults are proposed to be guided by a principle of parsimony which leads them to prefer weaker interpretations. In the context of our experiment, Crain et al. expected that that when accessing inference-based interpretations children would prefer interpretations including the None inference, while adults would prefer interpretations including the NotEvery inference. Our experiment produced the expected pattern of results; when participants accessed inference-based interpretations of our test sentences, children tended to compute the None inference, while adults tended to compute the NotEvery inference. This result lends support to the proposal, advanced by Crain et al., that adults and children are guided by different principles when interpreting sentences.

Finally, our results also provided an opportunity to adjudicate between competing accounts of scalar inferences. The Pragmatic account predicted that children would prefer the NotEvery inference, while the Grammatical account, assuming the subset principle, predicted that children would prefer the None inference. Therefore, our finding that children preferred interpretations including the None inference over the NotEvery inference provides further support for the Grammatical account as an explanation of how scalar inferences are derived.

4.4 Appendix

In this appendix, we show how a traditional scalar inference, the OnlySome inference is proposed to be derived. Following this we outline how the same process is used to account for the derivation of the NotEvery inference, and None inference. The Grammatical and Pragmatic accounts both derive the OnlySome and the NotEvery inference through similar processes (i.e. application of the scalar inference process at the whole-sentence level). However, they differ in how they capture the derivation of the None inference. Therefore, while we will outline only one derivation for the OnlySome and NotEvery inferences, we will provide each accounts' distinct derivation of the None inference.

4.4.1 Deriving the OnlySome inference

Asserted sentences like (55) have been associated with OnlySome inferences like (56).

(55) The pig carried some of his rocks.

(56) The pig didn't carry all of his rocks.

The OnlySome inference is proposed to be computed as a scalar inference through the steps laid out in (61a)-(61c).

(57) **Asserted sentence:** The pig carried some of his rocks.

(58) **Alternative sentences:** The pig carried all of his rocks.

(59) **Negated alternative sentences:** It's not the case that the pig carried all of his rocks.

(60) **Asserted sentence + Negated alternative sentences:** The pig carried some of his rocks AND It's not the case that the pig carried all of his rocks.

(61)
a. The asserted sentence in (57) is spoken, instead of the alternative sentences in (58).
b. The alternative sentences in (58) that are stronger than the asserted sentence in (57) are negated to generate the negated alternative sentences in (59).
c. If a negated alternative sentence in (59) does not contradict the asserted sentence in (57), it is combined with the asserted sentence in (57), thereby creating the final meaning in (60).

Next, we will show how the NotEvery inference can be derived as a scalar inference.

4.4.2 Deriving the NotEvery inference

Asserted sentences like (62) have been associated with NotEvery inferences like (63).

(62) Every pig carried some of his rocks.

(63) It's not the case that every pig carried all of his rocks.

The NotEvery inference is proposed to be computed as as a scalar inference through the steps laid out in (68a)-(68c).

(64) **Asserted sentence:** Every pig carried some of his rocks.

(65) **Alternative sentences:** Every pig carried all of his rocks.

(66) **Negated alternative sentences:** It's not the case that every pig carried all of his rocks.

(67) **Asserted sentence + Negated alternative sentences:** Every pig carried some of his rocks AND It's not the case that every pig carried all of his rocks.

(68) a. The asserted sentence in (64) is spoken, instead of the alternative sentences in (65).

b. The alternative sentences in (65) that are stronger than the asserted sentence in (64) are negated to generate the negated alternative sentences in (66).

c. If a negated alternative sentence in (66) does not contradict the asserted sentence in (64), it is combined with the asserted sentence in (64), thereby creating the final meaning in (67).

Finally, we will present the details of the None inference's computation as a scalar inference.

4.4.3 Deriving the None inference

Before presenting the detail of how the None inference is proposed to be computed as a scalar inference, we need to note a couple of things. First, the Pragmatic and Grammatical accounts of scalar inference computation differ in regards to how they capture the None inference. While versions of both accounts can explain the computation of None inferences from EverySome sentences as scalar inferences, the process through-which this is achieved is quite different.[13]

[13]Note that some versions of the Pragmatic account capture the None inference by appealing to alternative non-scalar inference mechanisms. We will not be detailing this process here. For more information about these proposals see Geurts and van Tiel (2013); van Tiel (2014).

The Grammatical account explains scalar inference computation by appealing to a covert grammatical operator, EXH. Such an operator is able, not only to be applied to whole-sentences, but can also be embedded within sentences. For example, EverySome sentences like (69) have two sites at which EXH can be applied; at the whole-sentence level, as in (70), or embedded under the universal quantifier, as in (71).

(69) Every pig carried some of his rocks.

(70) EXH[Every pig carried some of his rocks.]

(71) Every pig EXH[x carried some of his rocks.]

Applying EXH at the whole-sentence level results in the computation of the NotEvery inference in (72), as we just outlined (Section 4.4.2). However, embedding EXH under the universal quantifier results in the derivation of the None inference in (73). That is, applying EXH at this embedded location results in the scalar inference process being carried out on a sentence that is similar to the sentence we presented in the previous section outlining the derivation of the OnlySome inference (Section 4.4.1). Therefore, application of EXH to this part of (69) results in the derivation of an OnlySome inference like (74) through the process outlined in Section 4.4.1. And, when this OnlySome inference is combined with the universal quantifier, the inference in (75) is generated, which includes the None inference in (73).

(72) Not every pig carried all of his rocks.

(73) None of the pigs carried all of his rocks.

(74) [pig carried some but not all of his rocks.]

(75) Every pig carried some but not all of his rocks.

In contrast, the process through-which versions of the the Pragmatic account explain the computation of the None inference as a scalar inference is quite different (Spector, 2003; Sauerland, 2004; Chemla, 2009a). Unlike the Grammatical account, the Pragmatic account contends that it is not possible for the scalar inference process to be embedded within a sentence. Despite this, one way the Pragmatic account can explain the computation of the None inference is by adding the sentence in (76) to the alternative sentence set of the asserted sentence in (69).

As we have already mentioned, in order to include the sentence in (76) in this set of alternative sentences, two assumptions need to be added to the scalar inference process. First, it needs to be assumed that alternative sentences can be generated through the replacement of multiple lexical items. Second, it needs to assumed that alternative sentences can be logically independent of the relevant asserted sentence.

(76) Some pig carried all of his rocks.

The inclusion of (76) to the alternative sentences set of the assertion in (69) means that the None inference can now be derived through the application of the scalar inference process at the whole-sentence level. The steps through-which this is achieved are laid out in (81a)-(81c).

(77) **Asserted sentence:** Every pig carried some of his rocks.

(78) **Alternative sentences:** Some pig carried all of his rocks.

(79) **Negated alternative sentences:** It's not the case that some pig carried all of his rocks.

(80) **Asserted sentence + Negated alternative sentences:** Every pig carried some of his rocks AND It's not the case that some pig carried all of his rocks.

(81) a. The asserted sentence in (77) is spoken, instead of the alternative sentences in (78).

b. The alternative sentences in (78) that are stronger than the asserted sentence in (77) are negated to generate the negated alternative sentences in (79).

c. If a negated alternative sentence in (79) does not contradict the asserted sentence in (77), it is combined with the asserted sentence in (77), thereby creating the final meaning in (80).

Breaking down the final meaning in (80), there is the literal meaning in (82) and the inference in (83), which is equivalent to the None inference in (73).

(82) Every pig carried at least one of his rocks.

(83) It's not the case that some pig carried all of his rocks.

In sum, the NotEvery and None inferences, associated with EverySome sentences, can be captured as being computed through the scalar inference process by both the Grammatical account of scalar inferences, as well as versions of the Pragmatic account.

Chapter 5

Conclusion

5.1 Introduction

The present thesis investigated children's ability to compute a series of different *scalar inferences*. These phenomena all involve inferences that are derived from an *asserted sentence* through the negation of certain *alternative sentences*. Experimental investigations over the last few decades have revealed considerable variation in children's performance in computing scalar inferences (Noveck, 2001; Chierchia et al., 2001; Papafragou and Musolino, 2003; Barner et al., 2011; Tieu et al., 2016). This variation presents a challenge for theories attempting to explain children's pragmatic and grammatical abilities. One group of theories that have been quite successful in meeting this challenge is the *Alternatives-based approach*. This approach includes explanations that attribute children's variable success to limitations in processing capacity (Chierchia et al., 2001; Gualmini et al., 2001; Reinhart, 2006; Tieu et al., 2016) or to limitations in their knowledge of lexical scales (Barner and Bachrach, 2010; Barner et al., 2011). One prediction that can be drawn from the Alternatives-based approach is that, in cases where the computational requirements imposed on processing resources or on lexical access are lower, children are expected to derive scalar inferences readily. One situation where these requirements are conceivably reduced is when the linguistic material, from which the relevant alternative sentences are composed, is contained in the asserted sentence. In fact, many of the scalar inferences that children have been found to derive readily are associated with such sentences. In this way, the Alternatives-based approach is able to explain why children have often been found to struggle with the derivation of certain scalar inferences, as well as why, children succeed in the derivation of other scalar inferences.

In addition to explaining children's variable success in previous experimental studies of scalar inferences, the Alternatives-based approach makes specific predictions about how children will engage with certain scalar inferences that have not yet been investigated. According to the Alternatives-based approach, children are expected to be successful in computing scalar inferences for asserted sentences, when the alternatives can be composed from linguistic material contained in the asserted sentence. One way to test the Alternatives-based approach, therefore, is to investigate children's behaviour with such inferences, to see if it is consistent with this prediction. This thesis presented a series of such investigations.

Chapter 2 investigated children's derivation of *distributive inferences* and *conjunctive inferences*. Chapter 3 investigated *basic free choice inferences* and *universal free choice inferences*. Chapter 4 investigated two inferences that are associated with *EverySome sentences*; the *NotEvery inference* and the *None inference*. In general, the results of these investigations were in line with the predictions of the Alternatives-based approach, providing further support for this approach as a viable explanation of children's variable success in computing scalar inferences.

This thesis also contributed to an on-going debate between two contemporary theories about the mechanisms that are used in the derivation of scalar inferences. These are the *Pragmatic account* and the *Grammatical account*. These accounts appeal to different mechanisms to explain how people derive scalar inferences. The Pragmatic account is a descendant of the traditional account by Grice (1975) and Horn (1972). Like the traditional account, the mechanisms posited by the Pragmatic account are general reasoning and conversational norms (Sauerland, 2004; Geurts, 2010). The Grammatical account, by contrast, contends that scalar inferences are computed by using a grammatical operator, EXH (Chierchia, 2006; Fox, 2007; Chierchia, 2013). This grammatical operator is applied during the semantic compositional derivation of sentence meaning. On the Pragmatic account, the mechanisms invoked to compute scalar inference can only be applied to whole-sentences, whereas the EXH operator posited by the Grammatical account can apply to both whole-sentences and sentence internally. These differences in processing characteristics, in turn, yield differences in the classification of inferences, including the inferences that were investigated in the present thesis. The Grammatical account classifies all of the inferences we investigated as scalar inferences. The Pragmatic account, on the other hand, does not consider the universal free choice inference (see Chapter 3) to be a scalar inference. Rather, the Pragmatic account takes this inference to be derived through an alternative mechanism (Geurts and Pouscoulous, 2009b; Geurts, 2010). Moreover, the Pragmatic account views the None inference in Chapter 4 to be derived either through a more complex instantiation of the scalar inference process (Sauerland, 2004; Chemla, 2009a), or through an alternative mechanism (Geurts, 2010; Geurts and van Tiel, 2013; van Tiel, 2014). These differences in classification lead to different predictions about the behaviour that is expected from language users. The Grammatical account expects for users to produce a similar pattern of behaviour in generating all the inferences we investigated. More importantly, this pattern of behaviour is expected to mirror that of other scalar inferences. By contrast, the Pragmatic account could accommodate a wider range of behavioural outcomes.

In general, the results of our investigations found that the behaviour of adults and children in computing the inferences investigated was consistent with the Grammatical account. There was a continuity in adults and children's behaviour within the investigated inferences, and also between these inferences and other scalar inferences. As a result, we interpret our findings as providing additional support for the Grammatical account.

The rest of this chapter proceeds as follows. First, we present a summary of each of the investigations reported in this thesis. Based on the findings of these investigations, we then sketch the idea that recursively applying the scalar inference process is costly. Following this, we identify some promising directions for future research in this area. Finally, we conclude by summarising the main contributions of this thesis to the literature.

5.2 Chapter summary

This thesis investigated the interpretations children assigned to a series of sentences that have been associated with different scalar inferences. We now summarise the motivation and contribution of each of these investigations.

5.2.1 Children's distributive inferences

Chapter 2 investigated how children interpreted sentences in which the word for disjunction appears in the scope of the universal quantifier, as illustrated in (1). One inference that has been associated with sentences like (1) is the distributive inference indicated in (2) (Crnič et al., 2015). Another inference associated with assertions like (1), is the conjunctive inference in (3) (see, e.g., Singh et al. (2016)). This inference was also investigated in Chapter 2.

(1) Every elephant caught a big butterfly or a small butterfly.

(2) Some elephant caught a bug butterfly and some elephant caught a small butterfly.

(3) Every elephant caught both a big butterfly and a small butterfly.

For both the distributive inference and the conjunctive inference, it is possible to formulate the relevant alternatives using linguistic material contained in the asserted sentence in (1). According to Alternatives-based approach, therefore, the requirements for processing these inferences, including access to the relevant lexical scales, is lower for these inferences, than for some 'classical' scalar inferences. As a result, these inferences are expected to be easier for children to derive on the Alternatives-based approach. That is, the Alternatives-based approach expects for children to derive an inference-based interpretation (i.e. derive either the distributive inference or the conjunctive inference) of sentences like (1) at an adult-like rate. An experiment was designed to test this expectation by comparing the rates at which adults and children derived distributive inferences like (2) and conjunctive inferences like (3) from assertions like (1). We found that adults and children derived an inference-based interpretation of the target sentences at the same rate, as expected by the Alternatives-based approach.

Moreover, while the bulk of these inference-based interpretations were a result of having derived the distributive inference, there were some participants (mainly children, but even one adult) who generated a conjunctive inference. This aspect of children's behaviour is consistent with other recent work suggesting that, for children, the conjunctive inference should be a possible interpretation of sentences like (1) (Singh et al., 2016; Tieu et al., 2017).

5.2.2 Children's free choice inferences

Chapter 3 investigated children's interpretations of sentences where disjunction is in the scope of a deontic modal verb phrase, as in (4). Such assertions are associated with a basic free choice inference, as indicated in (5).

(4) The dog is allowed to carry a green stone or a purple stone.

(5) The dog is allowed to carry a green stone and the dog is allowed to carry a purple stone.

Chapter 3 also investigated children's interpretations of sentences where disjunction and a deontic modal are in the scope of a universal quantifier, as in (6). Assertions like this are associated with universal free choice inferences, as in (7).

(6) Every dog is allowed to carry a green stone or a purple stone.

(7) Every dog is such that it is allowed to carry a green stone and it is allowed to carry a purple stone.

Like the inferences investigated in Chapter 2, the assertions associated with basic free choice inferences and universal free choice inferences contain linguistic material from which their alternative sentences can be composed. Therefore, the Alternatives-based approach expects that, as the processing and lexical scale knowledge requirements for these inferences are reduced, children will derive them at an adult-like rate.

We tested this prediction by measuring the rates at which adults and children derived basic free choice inferences and universal free choice inferences. Within each participant group we found no difference in the rates at which these two free choice inferences were derived. Moreover, we found that children derived both free choice inferences at a higher rate than has been found with 'classical' scalar inferences (e.g., the exclusivity inference). However, the rate at which children derived both of these free choice inferences was still lower than the rates at which adults derived them. Considering this pattern of results, we concluded that while our results are in the direction expected by the Alternatives-based approach, the difference with adults requires some further explanation. One possibility we considered was that the recursive application of the scalar inference process is beyond the processing capabilities of some children. We will return to this idea momentarily.

The results of this investigation also contributed to the on-going debate between the Pragmatic and Grammatical accounts of scalar inferences. The Grammatical account categorises both variants of the free choice inference we investigated as scalar inferences (Kratzer and Shimoyama, 2002; Alonso-Ovalle, 2005; Fox, 2007). In contrast, the Pragmatic account can only categorise the basic free choice inference as a scalar inference. One strategy employed by the Pragmatic account is to adopt an alternative process to explain both kinds of free choice inference (Geurts and Pouscoulous, 2009b; Geurts, 2010). As a result, only the Grammatical account expects the patterns of behaviour for language users, in deriving free choice inferences to mirror that of other scalar inferences. Therefore, to the extent that children's behaviour with free choice inferences is consistent with the expectations of the Alternatives-based approach, our results provide support for the Grammatical account of scalar inference derivation. Moreover, the strong correlation we found within each participant group in regards to their performance between the two target inferences supports the idea that they are derived through the same process, whether that is the scalar inference process or some alternative.

5.2.3 Children's inferences from sentences with *every...some*

Chapter 4 investigated children's behaviour with two inferences associated with EverySome sentences like (8). The inferences licensed by (8) include the NotEvery inference in (9), and the None inference in (10).

(8) Every pig carried some of his rocks.

(9) Not every pig carried all of his rocks.

(10) None of the pigs carried all of his rocks.

Once again, the relevant asserted sentence contains the linguistic material required to compose the alternatives from which these inferences can be derived. In contrast to the assertions investigated in earlier chapters, however, the composition of the relevant asserted sentences require a greater level of reconstruction. That is, in the case of the distributive inferences investigated in Chapter 2, the relevant alternative sentences are presented as the asserted sentence's individual disjuncts. So, the alternatives can be generated simply by deleting parts of the asserted sentence. In contrast, while EverySome sentences contain the key lexical items, the alternative sentences are generated by replacing certain lexical items in the asserted sentence. For example, the alternative sentence in (11) is involved in the NotEvery inference's derivation. This alternative sentence is generated from the asserted sentence in (8) by replacing the existential quantifier 'some' with the universal quantifier 'all', a process that is conceivably assisted by the presentation in the asserted sentence of the universal quantifier 'every'.

(11) Every pig carried *all* of his rocks.

Chapter 4 explored whether providing the relevant linguistic material in this more incomplete manner would be sufficient for children to overcome the Alternatives-based approach's proposed limitations. This investigation involved measuring the rates at which adults and children accessed inference-based interpretations of EverySome sentences - that is, interpretations including the NotEvery inference in (9), or the None inference in (10).

As noted, the Alternatives-based approach did not predict which inference children should favour in generating inference-based interpretations for the asserted sentences we investigated. A relevant prediction was made, however, in earlier work on a different linguistic phenomenon, in Crain et al. (1994). These researchers proposed that adults and children are guided by different principles when interpreting sentences. Based on considerations of language learnability in the absence of negative evidence, young children are expected to prefer stronger (subset) interpretations. Initially selecting the stronger interpretation guarantees that children will have access to positive evidence - if adult speakers of the local language favour the weaker superset interpretations. This learnability principle is called the *subset principle* (Berwick, 1985). In contrast to children, Crain et al. propose that adults prefer weaker interpretations, so that they can avoid the cognitive costs associated with having to re-analyse sentences that are open to more than one interpretation. This parsing principle is called the *principle of parsimony* (Crain and Steedman, 1985). Considering EverySome sentences like (8), the principle of parsimony predicts that children will prefer the stronger None interpretation in (10) when they derive an inference-based interpretation, whereas adults will prefer the weaker NotEvery inference in (9).

Chapter 4 also contributed to the debate between the Pragmatic and Grammatical accounts of scalar inferences. These accounts differ in their expectations about the pattern of behaviour produced by children and adults in response to sentences like (8). As with the other inferences we investigated, the Grammatical account takes both these inferences to be derived through the application of the grammatical operator EXH to different parts of the assertion. In contrast, while the Pragmatic account can capture the NotEvery inference as a scalar inference, the None inference is more difficult for this account to explain. Some variants of the Pragmatic account (e.g., Sauerland (2004) and Chemla (2009a)), can capture the None inference through certain amendments to the scalar inference process. Specifically, these theories propose that the scalar inference process should allow the generation of alternative sentences through the replacement of multiple lexical items, and for alternative sentences to be logically independent of asserted sentences. Another way theories within the Pragmatic account have attempted to capture the None inference is as a non-conventional marginal interpretation generated through an alternative, non-scalar inference mechanism (Geurts, 2010; Geurts and van Tiel, 2013; van Tiel, 2014). Regardless of the strategy adopted, it is clear that there is a greater difference between the processes appealed to by the Pragmatic account to capture these inferences, compared to the processes proposed by Grammatical account. Moreover, the processes through which the Pragmatic account captures these inferences predicts that the NotEvery inference should be easier to access, and so preferred over the None inference. The previous experiments on adults' interpretations of EverySome sentences would appear to be consistent with this prediction, with adults preferring interpretations including the NotEvery inference over those including the None inference (Geurts and

Pouscoulous, 2009a; Chemla and Spector, 2011; van Tiel, 2014). However, as just mentioned, Crain et al. (1994) present an alternative explanation of this behaviour; that adults are guided by the principle of parsimony to prefer the weaker NotEvery interpretation. Investigating children's interpretations of EverySome sentences provided a unique opportunity to overcome this confound because, as mentioned, Crain et al. propose that children will prefer stronger interpretations. Therefore, if the Grammatical account is right, and these two inferences are derived through a very similar process, then, according to Crain et al., children should prefer interpretations including the None inference. On the other hand, if the Pragmatic account is right, and adults' previously documented behaviour is a result of the None inference being a marked reading or being derived through a more complex instantiation of the scalar inference process, then children are predicted to replicate adults in preferring interpretations including the NotEvery inference.

This chapter investigated the three theoretical issues just identified by exploring adults' and children's interpretations of EverySome sentences. Our experiment found that; a) adults and children derived inference-based interpretations at the same rate, and that b) the basis of these inference-based interpretations differed, with adults' preferring interpretations including the NotEvery inference, while children preferred those including the None inference. The finding that adults and children accessed inference-based interpretations at the same rate is consistent with the Alternatives-based approach. The finding that when deriving inference-based interpretations, adults preferred those including the NotEvery inference, whereas children preferred the None inference, is consistent with the proposal by Crain et al. (1994) - that adults prefer weaker interpretations, whereas children favour stronger interpretations. Finally, the observation that children derived the None inference readily, when accessing an inference-based interpretation is consistent with the Grammatical account of scalar inference derivation, over the Pragmatic account.

5.3 Processing cost and scalar inferences

Now we will present some ideas concerning the processing costs associated with scalar inference derivation. These ideas are inspired by a comparison of the results from the different investigations presented in the present thesis.

Chapter 3 investigated two kinds of free choice inference. According to the Grammatical account, both of these inferences are derived through the recursive application of the scalar inference process, although, this process is carried out at the embedded level when computing universal free choice inferences. Moreover, according to the Grammatical account, the None inference investigated in Chapter 4 is also derived by applying the scalar inference process at the embedded level. Assuming the Grammatical account's analysis, and considering the results of these investigations together provides evidence that it is costly to recursively apply the scalar inference process, whereas embedding the scalar inference process does not substantially increase the processing requirements. We now outline the reasoning that leads to this conclusion.

As mentioned earlier, an asserted sentence like (12) is associated with the basic free choice inference in (13). It has been proposed that this inference is derived through the recursive application of the scalar inference process, as indicated in (14) (Fox, 2007; Spector, 2007).

(12) The dog is allowed to carry a green stone or a purple stone.

(13) The dog is allowed to carry a green stone and the dog is allowed to carry a purple stone.

(14) EXH[EXH[The dog is allowed to carry a green stone or a purple stone.]]

Moreover, the sentence in (15) is associated with the universal free choice inference in (16), which, according to the Grammatical account, is derived by embedding this process under the universal quantifier, as indicated in (17).

(15) Every dog is allowed to carry a green stone or a purple stone.

(16) Every dog is allowed to carry a green stone and every dog is allowed to carry a purple stone.

(17) Every dog EXH[EXH[x is allowed to carry a green stone or a purple stone.]]

Chapter 3 reported no difference within each participant group, in the rates at which basic free choice inferences like (13) were derived, and the rates at which universal free choice inferences like (16) were derived. Moreover, there was a strong positive correlation within each participant group, in regards to their performance with each of these types of free choice inference. However, Chapter 3 reported a difference between participant groups in the rates at which both these free choice inferences were derived. As we mentioned in Chapter 3, one way to interpret this pattern of results is that, while embedding the scalar inference process does not increase the processing cost associated with scalar inference derivation, recursive application of the scalar inference process does increase the processing load. To put it another way, if we assume that; a) the Grammatical account is correct, and b) that an increase in an inference's processing costs will result in children deriving that inference at a lower rate than adults, then the following conclusions follow from our results. First, that there is no substantial increase in processing costs associated with embedding the scalar inference process, based on our finding of no difference in free choice inference derivation rates within each participant group. Second, that recursive application of the scalar inference process does involve a substantial processing cost, based on our finding that children derived both free choice inferences at a lower rate than adults.

These ideas, inspired by the results of Chapter 3, are further supported by the results of Chapter 4. Chapter 4 found that children's inference-based interpretations of EverySome sentences like (18) were largely based on having derived a None inference like (19). According to the Grammatical account, the None inference is derived through the application of the scalar inference process embedded under the universal quantifier, as in (20).

(18) Every pig carried some of his rocks.

(19) None of the pigs carried all of his rocks.

(20) Every pig EXH[x carried some of his rocks.]

We found that children were successful in deriving the None inference at the same rate as adults were deriving the NotEvery inference, which is thought to be derived through the scalar inference process being applied at the whole-sentence level. This result provides further support for the suggestion that embedding the scalar inference process does not substantially increase the processing cost of scalar inference derivation.

Finally, in Chapter 2, children were found to derive conjunctive inferences, which are derived through recursive application of the scalar inference process, at a much lower rate than distributive inferences, which are derived through a single application.[1] This further supports the suggestion that recursively applying the scalar inference process is costly.

We should note, however, that a study by Chemla and Bott (2014) produced results which appear, at least initially, to be inconsistent with this suggestion. Chemla et al. conducted a series of experiments designed to compare the processing costs associated with scalar inferences derived through the recursive application of the scalar inference process (i.e. free choice inferences), and a scalar inference derived through just one application of the scalar inference process, the *OnlySome inference* (i.e. some⇝not all). Chemla et al. did this by comparing the reaction times of responses associated with each of these inferences. A lower reaction time for a given response was interpreted as indicative of a lower processing cost for the associated inference. Chemla et al. found that the processing cost associated with the inferences derived through recursive application of the scalar inference process (i.e. free choice inferences) was lower than for inferences derived through a single application of the scalar inference process (i.e. OnlySome inferences). These results would appear to be inconsistent with the idea that recursive application of the scalar inference process is costly.

However, there is another possible interpretation of Chemla and Bott (2014)'s results, as Chemla et al. themselves identify. It could be that, as proposed by certain versions of the Alternatives-based approach (e.g., Chierchia et al. 2001; Reinhart 2006; Tieu et al. 2016), the processing requirements of deriving OnlySome inferences are already quite high, due to the nature of the alternative sentences through which they are derived. This observation is significant because, as we have already mentioned many times, the Alternatives-based approach takes the processing costs associated with deriving free choice inferences to be lower in this regard. Therefore, in the study by Chemla et al., the increase in processing load associated with the recursive application of the scalar inference process may have been masked by the relative processing costs associated with composing each inferences alternative sentences. To put it another way, it is possible that there are processing costs associated with recursively applying the scalar inference, but

[1] Although as we note in Chapter 2 (Section 2.3.2), elements of our experiment may have biased participants against accessing the conjunctive inference.

that there are even higher processing costs associated with generating the alternative sentences associated with OnlySome inferences. Given these considerations, it would appear that Chemla et al.'s results pose less of a problem for the suggestion that recursively applying the scalar inference process increases processing costs.

5.4 Future directions

Future work could proceed in a variety of directions. One avenue that has already produced useful results is the employment of on-line measures such as eye-tracking to provide further insight into the processing costs associated with deriving inferences. Such measures could be used to further test the Alternative-based approach's proposition that children's typical difficulties with scalar inference derivation are a result of limitations in children's processing capacity. Moreover, these measures could be used to explore further the suggestion that recursively applying the scalar inference process is associated with a processing cost.

Another promising avenue for research would be to investigate the inferences that are licensed by individuals who experience delays or deficits in language acquisition, such as children with Specific Language Impairment or Autism, or adults with aphasia. Such investigations could provide a further test of the Alternatives-based approach, as well as provide further insight into the specific linguistic abilities that are impaired in these populations.

Finally, it could also be useful to investigate how the inferences we investigated are expressed across different languages. This would be particularly interesting for languages that vary in regards to the linguistic material that is presented as part of the asserted sentence.

5.5 Conclusion

This thesis investigated children's behaviour with a series of different inferences that have a scalar inference analysis. The primary aim of these investigations was to test the explanatory power of the Alternatives-based approach; a group of explanations that attempt to capture children's variable success in scalar inference computation. In general, the results of these experiments were consistent with the expectations of the Alternatives-based approach. Therefore, our results provide further support for the Alternatives-based approach as a viable explanation of children's variable success in computing scalar inferences.

Moreover, the results of the investigations in Chapter 3 and 4 contributed to the on going debate between the Pragmatic and Grammatical accounts of scalar inference derivation. The Grammatical account was able to capture all the inferences we investigated by appealing to the same scalar inference process. In contrast, the Pragmatic account was only able to capture certain inferences by appealing to more complicated instantiations of the scalar inference process, or by invoking alternative mechanisms. Across our investigations, the computation of inferences by language users was consistent with the theoretical expectations of a unified approach scalar inferences. For this reason, we take our results to provide further support for the Grammatical account, rather than the Pragmatic account.

The results of the investigation into the NotEvery and None inferences in Chapter 4 also contributed to our understanding of how more general sentence interpretation strategies interact with inferences. That is, the results of this chapter support the subset principle - the proposal that children would prefer stronger interpretations. In addition, the results support the principle of parsimony - the proposal that adults prefer weaker interpretations.

Finally, considering the results of all our chapters together raises the possibility that, although the embedding of the scalar inference process does not incur a significant processing cost, the recursive application of the scalar inference process, does.

Overall, it is hoped that this thesis has increased our understanding of the nature of scalar inferences, the process through which children acquire and develop an understanding of scalar inferences, and how scalar inferences interact with general principles of sentence interpretation.

Bibliography

Alonso-Ovalle, L. (2005). Distributing the disjuncts over the modal space. In *Proceedings of the North East Linguistics Society (NELS)*, volume 35.

An, S. (2015). *The acquisition of constraints in child mandarin.* PhD thesis, Macquarie University.

Barker, C. (2010). Free choice permission as resource-sensitive reasoning. *Semantics and Pragmatics*, 3(10):1–38.

Barner, D. and Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive psychology*, 60(1):40–62.

Barner, D., Brooks, N., and Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition*, 118(1):84–93.

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

Berwick, R. (1985). *The Acquisition of Syntactic Knowledge.* Artificial Intelligence Series. MIT Press.

Bill, C., Romoli, J., Schwarz, F., and Crain, S. (2016). Scalar implicatures versus presuppositions: The view from acquisition. *Topoi*, 35(1):57–71.

Boster, C. and Crain, S. (1993). On children's understanding of every and or. In *Proceedings of early cognition and transition to language.*

Bowler, M. (2014). Conjunction and disjunction in a language without ‘and’. In *Semantics and Linguistic Theory*, volume 24, pages 137–155.

Breheny, R., Klinedinst, N., Romoli, J., and Sudo, Y. (2016). The symmetry problem: current theories and prospects. Under review, available at https://www.ucl.ac.uk/ ucjtudo/pdf/alternatives.pdf.

Chemla, E. (2009a). Similarity: Towards a unified account of scalar implicatures, free choice permission and presupposition projection. Unpublished, available at http://www.emmanuel.chemla.free.fr/Material/Chemla-SIandPres.pdf.

Chemla, E. (2009b). Universal implicatures and free choice effects: Experimental data. *Semantics and Pragmatics*, 2(2):1–33.

Chemla, E. and Bott, L. (2014). Processing inferences at the semantics/pragmatics frontier: disjunctions and free choice. *Cognition*, 130(3):380–396.

Chemla, E. and Spector, B. (2011). Experimental evidence for embedded scalar implicatures. *Journal of Semantics*, 28(3):359–400.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In *Structures and Beyond: The Cartography of Syntactic Structures*, volume 3, pages 39–103. Oxford University Press.

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the “logicality” of language. *Linguistic inquiry*, 37(4):535–590.

Chierchia, G. (2013). *Logic in Grammar: Polarity, Free Choice, and Intervention.* Oxford University Press.

Chierchia, G. (2017). Scalar implicatures and their interface with grammar. *Annual Review of Linguistics*, 3(1):245–264.

Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A., and Meroni, L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In *Proceedings of the 25th Boston University conference on language development*, pages 157–168. Cascadilla Press Somerville, MA.

Chierchia, G., Fox, D., and Spector, B. (2011). The grammatical view of scalar implicatures and the relationship between syntax and semantics. In Maienborn, C., von Heusinger, K., and Portner, P., editors, *Semantics: An International Handbook of Natural Language Meaning*, volume 1. Mouton de Gruyter, New York.

Choi, Y. and Mazuka, R. (2003). Young children's use of prosody in sentence parsing. *Journal of Psycholinguistic Research*, 32(2):197–217.

Clifton Jr, C. and Dube, C. (2010). Embedded implicatures observed: A comment on Geurts and Pouscoulous (2009). *Semantics and Pragmatics*, 3(7):1–13.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, volume 2. Hillsdale, NJ: Earlbaum.

Crain, S. (1992). The semantic subset principle in the acquisition of quantification. In *Workshop on the Acquisition of WH-Extraction and Related Work on Quantification, University of Massachusetts, Amherst, MA*.

Crain, S., Ni, W., and Conway, L. (1994). Learning, parsing and modularity. In Clifton Jr, C., Frazier, L., and Rayner, K., editors, *Perspectives on sentence processing*, pages 443–467. Lawrence Erlbaum Associates Hillsdale, NJ.

Crain, S. and Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. In Dowty, D., Kartunnen, R., and Zwicky, A., editors, *Natural language parsing*, pages 320–358. Cambridge University Press, Cambridge.

Crain, S. and Thornton, R. (1998). *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics.* Language, speech, and communication. MIT Press, Cambridge.

Crnič, L., Chemla, E., and Fox, D. (2015). Scalar implicatures of embedded disjunction. *Natural Language Semantics*, 23(4):271–305.

Davidson, K. (2013). 'and' or 'or': General use coordination in asl. *Semantics and Pragmatics*, 6(4):1–44.

Foppolo, F., Guasti, M. T., and Chierchia, G. (2012). Scalar implicatures in child language: Give children a chance. *Language learning and development*, 8(4):365–394.

Fox, D. (2007). Free choice and the theory of scalar implicatures. In Sauerland, U. and Stateva, P., editors, *Presupposition and Implicature in Compositional Semantics*, pages 71–120. Palgrave Macmillan UK, London.

Fox, D. and Katzir, R. (2011). On the characterization of alternatives. *Natural Language Semantics*, 19(1):87–107.

Fuson, K. C. (1988). *Children's counting and concepts of number*. Springer-Verlag, New York.

Gamut, L. (1991). *Logic, Language, and Meaning, volume 1: Introduction to Logic*, volume 1. University of Chicago Press, Chicago.

Gazdar, G. (1979). *Pragmatics: Implicature, Presupposition, and Logical Form*. Academic Press, New York.

Geurts, B. (2005). Entertaining alternatives: Disjunctions as modals. *Natural language semantics*, 13(4):383–410.

Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press, Cambridge.

Geurts, B. and Pouscoulous, N. (2009a). Embedded implicatures?!? *Semantics and Pragmatics*, 2(4):1–34.

Geurts, B. and Pouscoulous, N. (2009b). Free choice for all: a response to Emmanuel Chemla. *Semantics and Pragmatics*, 2(5):1–10.

Geurts, B. and van Tiel, B. (2013). Embedded scalars. *Semantics and Pragmatics*, 6(9):1–37.

Grice, H. P. (1975). Logic and conversation. In Cole, P. and L, M. J., editors, *Syntax and semantics*, pages 41–58. Academic Press, New York.

Gualmini, A., Crain, S., Meroni, L., Chierchia, G., and Guasti, M. T. (2001). At the semantics/pragmatics interface in child language. In *Semantics and Linguistic Theory*, volume 11, pages 231–247.

Guasti, M. T., Chierchia, G., Crain, S., Foppolo, F., Gualmini, A., and Meroni, L. (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and cognitive processes*, 20(5):667–696.

Hochstein, L., Bale, A., Fox, D., and Barner, D. (2016). Ignorance and inference: do problems with gricean epistemic reasoning explain children's difficulty with scalar implicature? *Journal of Semantics*, 33(1):107–135.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70.

Horn, L. R. (1972). *On the semantic properties of logical operators in English*. PhD thesis, University of California, Los Angeles.

Huang, A. and Crain, S. (2014). Acquisition of the polarity sensitive item renhe 'any'in mandarin chinese. *Journal of child language*, 41(4):861–889.

Huang, Y. T. and Snedeker, J. (2009). Semantic meaning and pragmatic interpretation in 5-year-olds: Evidence from real-time spoken language comprehension. *Developmental Psychology*, 45(6):1723–1739.

Huang, Y. T., Spelke, E., and Snedeker, J. (2013). What exactly do numbers mean? *Language Learning and Development*, 9(2):105–129.

Jaeger, T. F. (2008). Categorical data analysis: Away from anovas (transformation or not) and towards logit mixed models. *Journal of memory and language*, 59(4):434–446.

Katsos, N. and Bishop, D. V. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, 120(1):67–81.

Katzir, R. (2007). Structurally-defined alternatives. *Linguistics and Philosophy*, 30(6):669–690.

Kiguchi, H. and Thornton, R. (2016). Connectivity effects in pseudoclefts in child language. *Studia linguistica*, 170(1):34–65.

Klinedinst, N. W. (2007). *Plurality and possibility.* PhD thesis, University of California, Los Angeles.

Kratzer, A. and Shimoyama, J. (2002). Indeterminate pronouns: The view from japanese. In *Proceedings of the Third Tokyo Conference on Psycholinguistics*, pages 1–25.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature.* MIT Press, Cambridge.

Magri, G. (2009). A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural language semantics*, 17(3):245–297.

Meyer, M.-C. (2013). *Ignorance and grammar.* PhD thesis, Massachusetts Institute of Technology.

Moscati, V., Romoli, J., Demarie, T. F., Crain, S., et al. (2016). Born in the usa: a comparison of modals and nominal quantifiers in child language. *Natural Language Semantics*, 24(1):79–115.

Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, 78(2):165–188.

Papafragou, A. and Musolino, J. (2003). Scalar implicatures: experiments at the semantics–pragmatics interface. *Cognition*, 86(3):253–282.

Philip, W. (1991). Quantification over events in early universal quantification. In *16th annual Boston University Conference on Language Development, Boston, MA.*

Philip, W. (1995). *Event quantification in the acquisition of universal quantification.* PhD thesis, University of Massachusetts.

Recanati, F. (2003). Embedded implicatures. *Philosophical perspectives*, 17(1):299–332.

Reinhart, T. (2006). *Interface Strategies. Optimal and Costly Derivations.* MIT Press, Cambridge.

Romoli, J. (2012). *Soft but strong. Neg-raising, soft triggers, and exhaustification.* PhD thesis, Harvard University.

Romoli, J. (2013). A scalar implicature-based approach to neg-raising. *Linguistics and philosophy*, 36(4):291–353.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and philosophy*, 27(3):367–391.

Schulz, K. (2006). A pragmatic solution for the paradox of free choice permission. In *Uncertainty, Rationality, and Agency*, pages 289–323. Springer Netherlands, Dordrecht.

Simons, M. (2005). Dividing things up: The semantics of or and the modal/or interaction. *Natural Language Semantics*, 13(3):271–316.

Singh, R., Wexler, K., Astle-Rahim, A., Kamawar, D., and Fox, D. (2016). Children interpret disjunction as conjunction: Consequences for theories of implicature and child development. *Natural Language Semantics*, 24(4):305–352.

Skordos, D. and Papafragou, A. (2016). Children's derivation of scalar implicatures: Alternatives and relevance. *Cognition*, 153:6 – 18.

Snedeker, J. and 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.

Spector, B. (2003). Scalar implicatures: Exhaustivity and gricean reasoning. In *Proceedings of the European Summer School in Logic, Language, and Information (ESSLLI) 15*, volume 3, pages 277–288.

Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures. In Sauerland, U. and Stateva, P., editors, *Presupposition and Implicature in Compositional Semantics*, pages 243–281. Palgrave Macmillan UK, London.

Stiller, A. J., Goodman, N. D., and Frank, M. C. (2015). Ad-hoc implicature in preschool children. *Language Learning and Development*, 11(2):176–190.

Thomas, G. G. P. Y. (2012). *Temporal implicatures.* PhD thesis, Massachusetts Institute of Technology.

Tieu, L., Bill, C., Romoli, J., and Crain, S. (2014). Plurality inferences are scalar implicatures: Evidence from acquisition. In *Semantics and Linguistic Theory*, volume 24, pages 122–136.

Tieu, L., Romoli, J., Zhou, P., and Crain, S. (2016). Children's knowledge of free choice inferences and scalar implicatures. *Journal of Semantics*, 33(2):269–298.

Tieu, L., Yatsushiro, K., Cremers, A., Romoli, J., Sauerland, U., and Chemla, E. (2017). On the role of alternatives in the acquisition of simple and complex disjunctions in french and japanese. *Journal of Semantics*, 34(1):127–152.

Trinh, T. and Haida, A. (2015). Constraining the derivation of alternatives. *Natural Language Semantics*, 23(4):249–270.

van Tiel, B. (2014). Embedded scalars and typicality. *Journal of Semantics*, 31(2):147–177.

Zhou, P., Romoli, J., and Crain, S. (2013). Children's knowledge of free choice inferences. In *Semantics and Linguistic Theory*, volume 23, pages 632–651.

Zhou, P., Su, Y. E., Crain, S., Gao, L., and Zhan, L. (2012). Children's use of phonological information in ambiguity resolution: a view from mandarin chinese. *Journal of Child Language*, 39(4):687–730.

Zimmermann, T. E. (2000). Free choice disjunction and epistemic possibility. *Natural language semantics*, 8(4):255–290.