Disjunction triggers exhaustivity implicatures in 4- to 5-year-olds: Investigating the role of access to alternatives

Nicole Gotzner[a], David Barner[b] and Stephen Crain[c]

a Zentrum für Allgemeine Sprachwissenschaft, b UCSD, c Macquarie University

Short title: Exhaustivity and disjunction

Address for Correspondence:

Nicole Gotzner

Zentrum für Allgemeine Sprachwissenschaft

Schützenstr.18

10117 Berlin

Germany

gotzner@zas.gwz-berlin.de

Abstract

Children's difficulty deriving scalar implicatures has been attributed to a variety of factors including processing limitations, an inability to access scalar alternatives, and pragmatic tolerance. The present research explores the nature of children's difficulty by investigating a previously unexplored kind of inference – an exhaustivity implicature that is triggered by disjunction. We reasoned that if children are able to draw quantity implicatures but have difficulties accessing alternative lexical expressions from a scale, then they should perform better on exhaustivity implicatures than on scalar implicatures, since the former do not require spontaneously accessing relevant scalar alternatives from the lexicon. We conducted two experiments. Experiment 1 found that 4- to 5-year-olds consistently computed exhaustivity implicatures to a greater extent than scalar implicatures. Experiment 2 demonstrated that children are more likely to compute exhaustivity implicatures with disjunction compared to conjunction. We conclude that children often fail to derive scalar implicatures because (1) they struggle to access scalar alternatives and (2) disjunction (but not conjunction) makes subdomain alternatives particularly salient. Thus, the findings suggest that exhaustivity implicatures can be derived without reference to a scale of alternatives.

Keywords: scalar implicatures, pragmatic inferences, exhaustivity, disjunction, child language acquisition

Introduction

To become competent speakers of a natural language, children must acquire a compositional semantic system that supports the creation of indefinitely many meanings from a finite inventory of semantic representations. Moreover, the literal meanings of the expressions contained in the lexicon of the semantic system do not alone explain the expressive power of natural languages: Literal meanings are often strengthened contextually. To strengthen expressions, listeners add to the literal content of sentences by taking into consideration alternative utterances that a speaker might have used, but elected not to use. Consider, for example, the sentences in (1) and (2):

(1) Some of the animals are sleeping.

(2) All of the animals are sleeping.

In a context in which a speaker utters sentence (1), the listener infers that the speaker believes sentence (2) to be false, since if the speaker were in a position to know that (2) was true, then she would have said so (in accordance with the conventional norm to be as informative as possible). This type of inference, which involves an implicit comparison of ordered statements, is called a scalar implicature. The utterance that is actually produced is compared to alternatives that are generated via the replacement of the scalar item *some* with *all*, which is selected from a scale that orders expressions according to the strength of the information they convey. If a speaker uses an expression that is weaker than some other expression on the scale, then the hearer should infer that the speaker was not in a position to use a stronger expression, and thereby negate the stronger expression. This results in a derived expression – e.g., *Some, but not all, of the animals are sleeping.*

From the perspective of acquisition, the distinction between basic and derived meanings could be problematic because of the inductive problem faced by child language learners. Given the distinction between basic and derived meanings, how do language learners come to know which aspects of meaning are computed in the semantic component and which are computed in the pragmatic component? At least since Grice (1975), a similar problem has confronted linguists, who are also faced with the problem of differentiating information that is computed at the semantic and pragmatic levels of representation. One recent solution to this problem contends that implicatures are computed on-line, as part of the semantic system (e.g., Chierchia, 2006; Chierchia, Fox & Spector, 2012), or they are computed by default without appealing to pragmatics (Levinson, 2000). In contrast, advocates of Gricean models locate only basic meanings in the semantic component, and attribute implicatures to independent pragmatic principles such as the Maxim of Quantity (e.g., Grice, 1975; Gazdar, 1979; Sauerland, 2004).

Empirical studies of language acquisition afford the opportunity to adjudicate between these alternative models and to address how children differentiate semantics and pragmatics. This can be achieved by asking which factors in development best predict changes in children's ability to compute implicatures. If implicatures are essentially a form of epistemic enrichment that is overlaid onto semantic representations, as Grice and his theoretical descendants contend, then changes in children's ability to compute implicatures may be related to the development of theory of mind – the ability to reason about the beliefs of interlocutors. On the other hand, if implicatures involve purely semantic resources then children's abilities to compute implicatures may develop somewhat independent of theory of mind and instead be related to the development of grammatical knowledge.

Although a large body of studies has documented children's difficulties computing implicatures, there remains significant disagreement regarding the nature of their behaviors. In a range of different experimental paradigms, children differ from adults, and accept sentences that contain weak scalar terms like (1), when both (1) and the stronger statements in (2) are also true (e.g., Chierchia et al., 2001; Gualmini et al., 2001; Huang & Snedeker, 2009a; Noveck, 2001; Papafragou & Musolino, 2003; Guasti et al., 2005; Barner, Chow, & Yang, 2009; Foppolo et al., 2012). An early study by Noveck (2001) reported that children up to age 11 have difficulties in deriving scalar implicatures with modal verbs (*might* and *must*), and with quantificational expressions like *some.* Other research has reported that children are relatively insensitive to exclusivity inference associated with disjunction (Paris, 1973; Smith, 1980 and Braine & Rumain, 1980). As another example, using a felicity judgment task, Papafragou and Musolino (2003) found that 5-year-old children struggle to compute implicatures involving scales like *finish/start*. The majority of these studies included adult control groups, who consistently computed implicatures to a greater degree than children.

According to some researchers, children's poor performance in computing implicatures provides evidence for Gricean accounts of conversational implicatures. Children appear to be more logical than adults (Noveck, 2001), on this view, because they are less sensitive than adults to certain communicative principles that speakers and hearers adhere to when they are engaged in conversation. One possible account attributes the observed differences between children and adults to children's developing theory of mind (see Flavell, 2004, for a review; Papafragou & Musolino, 2003).

A more recent account attributes the observed differences to children's greater tolerance to violations of pragmatic principles. Evidence in favor of the "pragmatic tolerance" hypothesis was reported in Katsos and Bishop (2011). These researchers demonstrated that the same 5- to 6-year-old children who accept under-informative utterances in a standard binary (right or wrong) judgment task, nevertheless differentiated between less informative statements and more informative statements when they were presented with a reward system that included an 'lesser' reward for speakers who produced true, but less than optimal descriptions of the situation. While Katsos and Bishop conclude from this that children are sensitive to violations of informativeness but judge them less harshly than adults, their data are also compatible with children's uncertainty about underinformative statements, as compared to statements that are maximally informative.

Others have argued that children's lack of sensitivity to certain inferences are due to processing limitations (see Chierchia et al., 2001; Pouscoulous et al., 2007; Reinhart, 2004). This conclusion is supported by the finding, from a number of studies, that children's performance in computing scalar inferences varies depending on the task demands (e.g., Papafragou & Musolino 2003; Papafragou & Tantalou 2004; and Pouscoulous et al. 2007). It remains unclear, however, whether task demands alone can account for the large differences in performance between children and adults (DeNeys & Schaeken, 2007; Antoniou, Cummins & Katsos, 2016). It is not known, for example, whether individual differences in computing implicatures are directly related to differences in processing capacity. While many studies have shown that adults take longer to compute implicatures than to compute literal meanings (e.g., Noveck & Posada, 2003; Huang & Snedeker, 2009b), it is unclear whether this increase in processing time is related to the developmental delay experienced by children in computing pragmatic inferences.

An additional account, and one that is not necessarily incompatible with the first two, is that children struggle to compute implicatures because they have difficulty accessing the relevant scalar alternatives from their mental lexicons (see Barner & Bachrach, 2010; Foppolo, et al., 2012; Skordos & Papafragou, 2014; Stiller, Goodman & Frank, 2015). Upon hearing a sentence with a weak scalar term, such as (1), children may not readily access the alternative sentence that replaces the weak scalar term with a stronger term, as in example (2). This may be due to a developmental delay in children's capacity to encode the expressions that are members of the relevant scale (which are known to differ cross-linguistically and, thus, must be learned). Alternatively, the various members of scales could be in place, but only weakly associated with one another in the child's mental lexicon. On this account, deficits in children's verbal working memory may be relevant in explaining their difficulties in computing scalar implicatures.

Several findings support this "access to alternatives" account. First, nearly every study that has tested children on multiple scalar contrasts has found differences across scales. Children as young as 2-years-old compute implicature-like inferences for numerals (Wynn, 1992), despite failing to compute scalar inferences with in sentences with *some* (versus *all*) and *or* (versus *and*) until much later (see Barner & Bachrach, 2010).[1] Similarly, older children are better with *some/all* than with *start/finish* and *might/must* (e.g., Papafragou & Musolino, 2003). Second, children readily identify the more informative of two utterances containing scalar items when the alternatives are made explicit on the same trial (e.g., Chierchia et al., 2001). Third, and perhaps most significantly, several studies have found that children perform significantly better on scalar inference tasks when alternatives are provided on other trials in the experiment. In one study, Skordos and Papafragou (2014) showed that children are significantly more likely to compute *some/all* implicatures if test trials containing *some* (e.g., *Some of the blickets have shovels*) are preceded by test trials containing *all* (see also Foppolo et al., 2012). In another study, Barner, Brooks, and Bale (2011) showed that access to alternatives affects how 4-year-old children interpret sentences even when the utterances *entail* the negation of stronger alternatives, e.g., when the focus particle *only* is used in questions like those in (3) and (4), in a context in which three animals were sleeping (a cat, a cow, and a dog).

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

(4) Are only some the animals sleeping?

Barner et al. found that, whereas both 4-year-olds and adults answered "No" to (3), only adults rejected (4) with the focus particle *only*. To explain children's affirmative answers to questions like (4), Barner et al. proposed that in interpreting sentences such as (4) children fail to spontaneously access the stronger utterance containing the lexical item *all* as a relevant alternative to *some.* In sentences such as (3), on the other hand, the relevant alternatives are introduced contextually, precluding the need for children to access stored lexical representations. Based on this, they argued that children's ability to compute *some/all* implicatures might be similarly restricted by access to alternatives.

---

[1] Yet an alternative account of this different behavior of numerals is that exact readings are part of their semantics.

A number of additional findings are consistent with the "access to alternatives" account. First, as noted earlier, children have been found to compute implicature-like inferences as young as 3 years of age when they are presented with the relevant alternatives in a forced choice task (Barner, Hochstein, Rubenson, & Bale, 2018; Chierchia et al., 2001; Stiller et al., 2015; Miller, Chang, Schmitt, & Munn 2005). In these studies, children's performance improved when they were provided either with the relevant linguistic alternatives, or when the visual scenes depicted referents that could be described by relevant linguistic alternatives. These forced-choice inferences have been found to be especially easy for children when they involve contextually defined *ad hoc* scales, rather than lexical scales like *some/all* (Stiller et al., 2015; Horowitz & Frank, 2015).

Especially relevant to the present study, several recent reports have found that access to alternatives limits children's ability to compute implicatures for disjunction. In one recent study by Tieu et al. (2016), children successfully computed free-choice inferences when they were presented with statements such as "Kung Fu Panda may push the green car or the orange car." (see 5 below) That is, children in the Tieu et al. study interpreted this sentence to imply that the Kung Fu Panda had been given permission to push both cars. However, the same children struggled to compute exclusivity implicatures (as in Braine & Rumain, 1981; Chierchia, et al., 2001; Paris, 1973). According to Tieu et al., children derived free-choice inferences because the relevant alternatives could be constructed (as in 6 and 7) from elements present in the original utterance – the subdomain alternatives contained within the disjunctive statement. According to their hypothesis, children lack access to the stronger alternative containing *and* (in 8), and therefore are unable to compute a scalar implicature (since negating the subdomain alternatives alone would result in contradicting the original utterance).

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

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

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

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

Following a similar logic, Hochstein, Bale, Fox, and Barner (2016) showed that whereas 5-year-old children fail to compute scalar implicatures involving disjunction, they readily compute ignorance implicatures. Like free-choice inferences, ignorance implicatures only involve subdomain alternatives (e.g., “The boy ate pizza”; “The boy ate pasta”). In this case, an utterance like “The boy ate pizza or pasta for lunch” (9) implies that the speaker does not believe that the boy ate pizza (resulting in the inference in 10), and also that they do not believe that the boy ate pasta (resulting in the inference in 11). Together, this leads to the joint inference that the speaker believes the original utterance – that one of these two foods was eaten – but simply is ignorant about which is true. By demonstrating that 5-year-olds can compute ignorance implicatures but not scalar implicatures, Hochstein et al. ruled out various deflationary accounts of children’s difficulties with implicature, including those that posit tolerance, processing limits, etc., since children were both willing and able to compute implicatures when alternatives were contextually available. Also, this study showed that children are competent Gricean reasoners and that problems with theory of mind cannot easily explain their difficulties in computing implicatures.

(9) B (The boy ate pizza or pasta for lunch).

(10) ¬B (The boy ate pizza for lunch).

(11) ¬B (The boy ate pasta for lunch).

Other recent work has also investigated the case study of disjunction to explore the “access to alternatives” hypothesis with some arguing that, on certain models of exhaustification (in particular Fox, 2007), failure to access a conjunctive alternative should lead children to make wildly non-adult-like inferences, including the inference that disjunctive statements imply conjunction. Specifically, when exhaustification is applied recursively to an utterance like the one in (12) such that subdomain alternatives –e.g., pizza and pasta – are first strengthened, negating the universally quantified disjuncts (13 and 14) results in the inference in (15), that the boy ate both pizza and pasta. While such an inference would be impossible given access to the stronger conjunctive scale mate, it is predicted if children lack access to the alternative including *and*. Based on this logic, studies first by Singh et al. (2016) and then by Tieu et al. (2016) argued that children often derive conjunctive interpretations of disjunctive statements, though others have found that children do not derive such inferences (Skordos, Feiman, Bale, & Barner, under review).

(12) Every boy ate pizza or pasta for lunch.

(13) ¬Every boy ate (only) pasta for lunch.

(14) ¬Every boy ate (only) pizza for lunch.

(15) Every boy ate both pizza and pasta for lunch.

To summarize the present state of affairs, previous studies have found that 4- to 5-year-old children experience difficulty computing implicatures triggered by scalar items such as *some* and *or,* but that children as young as 3 are able to derive implicatures when the relevant alternatives are either contextually salient or are presented in the linguistic context. Despite the large number of studies in the literature, however, there is no conclusive account of the differences between children and adults. Although Barner et al. (2011) showed that children exhibited differential sensitivity to both *ad hoc* and lexical scales like *some/all*, children's success was limited to questions containing the focus particle *only*. In the case of explicit *only* alternatives are negated as part of the lexical semantics of the focus particle and not by implicature. It remains possible, therefore, that although children are able to access the lexical expressions contained in contextually-salient scales, they are unable to compute *ad hoc* implicatures under similar circumstances. In the studies where children were most successful in responding to quantity implicatures, either the linguistic alternatives were provided to children or the alternative meanings were made highly salient in the nonlinguistic context (e.g., Barner et al., 2018; Chierchia et al., 2001; Miller et al., 2005; Katsos & Bishop, 2011; Stiller et al., 2015).

The present study: goals and hypotheses

The present study investigated children's ability to compute a form of *ad hoc* implicature that has not been studied previously. This enabled us to evaluate children's ability to access alternatives in a novel way. The experimental sentences introduced a kind of *ad hoc* implicature, which is referred to as an exhaustivity implicature in formal work on focus and implicature (e.g., van Rooij & Schulz, 2004; Spector, 2006). More specifically, we investigated the exhaustivity implicature that is derived from disjunctive statements. Consider a context in which there is a tiger, a penguin, and a pig. The research question is how children interpret the statement in (16) as a prediction about which of the animals will have a ball. This is called the *prediction mode* of a Truth Value Judgment task (Crain and Thornton, 1998).

(16) The tiger or the penguin will have a ball.

(17) The penguin or the tiger will have a ball, but not both.

(18) The penguin or the tiger, but not the pig, will have a ball.

Intuitively, the statement in (16) invites the inference that, as the situation unfolds, either the tiger or the penguin will have a ball, but that the speaker does not know which of them will have a ball (see Sauerland, 2004 for discussion, and see Hochstein et al., 2016 for evidence that 4- to 5-year-olds compute ignorance inferences). Sentence (16) also implicates (17), i.e., that not both animals will have a ball. So, sentence (16) licenses a scalar implicature, based on a comparison of sentence (16) with stronger alternative statements involving only one of individuals referred to in the disjunctive phrase (*The penguin will have a ball* and *The tiger will have a ball*) as well as the conjoined statement (*The tiger and the penguin will have a ball*). Each of these statements is stronger than the assertion made by (16) (see Sauerland, 2004; Fox, 2007). Finally, and critical to the present study, sentence (16) also generates the implicature in (18) - that no other contextually relevant animal will have a ball. That is, in the process of negating stronger alternatives, other contextual alternatives (i.e., the pig) can be added (see Barner et al., 2011).

As suggested above, many accounts of scalar inference propose that the scalar implicature in (17) is due to competition between *or* and its more informative scale-mate *and*. There are alternative accounts, however, that derive scalar inferences through the exhaustification of subdomain alternatives (e.g., van Rooij & Schulz, 2004; Schulz & van Rooij, 2006; Spector, 2006). This mechanism for exhaustification formalizes Neo-Gricean reasoning underlying the exclusivity inference without making essential reference to the scalar element *and* (Spector, 2016:3 for a comparison of different exhaustivity operators). Crucially, as van Rooij and Schulz (2004:501) point out, in examples like (19) exhaustive interpretation immediately predicts that only one of the disjuncts is true.

(19) (A: Who knows the answer?)

B: Peter, Mary, or Sue.

If children have the pragmatic capacity to derive implicatures but fail to spontaneously access scalar alternatives from the lexicon, then they should derive the exhaustivity inference in (18), even when they fail to derive the scalar implicature in (17). Specifically, this exhaustivity inference could be computed by negating the simple statement *The pig has a ball*[2] – i.e., if we assume that they have access to both the subdomain alternatives, e.g., in a statement of the form S(A or B), the individual alternatives S(A), S(B) as well as the salient contextual alternatives like S(C).

In the present study, we investigated the role of access to alternatives in scalar implicature by testing children with exhaustivity implicatures such as the one triggered by (16),

[2] It is an open question whether listeners build *ad hoc* scales/partially-ordered sets (e.g., Hirschberg, 1985) in these cases or they are guided by considerations or relevance, essentially deriving this a “nothing else” inference (Geurts, 2010). We will return to this question in the general discussion.

in situations where the relevant subdomain alternatives were made contextually salient. Four- to 5-year-old children were presented with scenes containing three different characters (a tiger, a pig, and a penguin) and a puppet who watched the scenes unfold alongside the child participant made predictions about the future outcome of each scene– e.g., "I think the tiger or the penguin will have a ball". In the final outcome (see Figure 1), one of the two disjuncts was confirmed – e.g., the tiger had a ball – but critically, an unmentioned alternative (e.g., the pig) also had a ball.

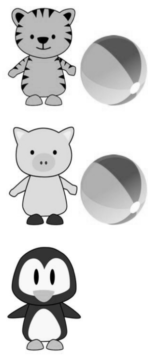


Figure 1: Example stimulus (outcome scene). Test sentence: "(I think) the tiger or the penguin will have a ball".

We predicted that if 4- to 5-year-old children derive exhaustivity implicatures, then they should reject the puppet's prediction and justify their rejections on the grounds that the third animal (the pig) had the same property as one of the animals mentioned in the disjunction phrase.

The second component of this study was designed to directly assess children's access to alternatives. This was achieved by comparing children's computation of *exhaustivity* implicatures and scalar implicatures. For this comparison we chose the scalar implicature triggered by the existential expression *some*, as in in Experiment 1. For adults, a speaker's use of a sentence with the existential expression *some* often implicates the negation of the corresponding sentences in which *some* has been replaced by a stronger alternative, e.g., *all.* Using the prediction mode paradigm of the Truth Value Judgment task, children were presented with statements like, "I think some of the animals will have a ball" preceding an outcome in which all of the depicted animals had a ball. We chose to examine the weak quantifier *some* in order to directly follow up on the Barner et al (2011) study, and because *some* has received considerable attention in the previous literature. In Experiment 1, children were presented with either a disjunctive statement or a statement involving the existential expression *some* using a between-subjects design. To foreshadow the outcome, we found that children were much more likely to compute an exhaustivity inference for disjunctive statements than *the scalar inference that some implicates not all.*

Experiment 2 compared children's generation of exhaustivity implicatures in disjunctive and conjunctive statements, in addition to using the scalar implicature of disjunction as a control. Conjunctive sentences like *The tiger and the penguin will have a ball* also have S(A) and S(B) as subdomain alternatives but these are not informationally stronger than the original statement and neither is the additional contextual alternative S(C).[3] That is, the only informationally stronger alternative would be an *ad hoc* scale involving all three alternatives S(A and B and C) but this is of higher complexity. Under the assumption of the structural account of alternatives by Fox and Katzir (2011) alternatives should be at most as complex as the prejacent, hence the alternative S(A and B and C) will not be derived. Participants should therefore focus on checking whether the two mentioned individuals have the property when presented with conjunctive statements. Since this situation is verified in the outcome scene, participants should accept the statement.

Hence, the experimental hypothesis of Experiment 2 was that children would be more likely to compute exhaustivity implicatures with disjunction than with conjunction. Again, we

[3] Further, an *ad hoc* scale involving the additional contextual alternative S(A and B and C) is of higher complexity than the original statement.

predicted that the child participants would fail to derive scalar implicatures for sentences with disjunction. As we will see, these predictions were borne out in Experiment 2.

Experiment 1

Experiment 1 was designed to compare two different forms of implicature: an exhaustivity implicature triggered by the logical connective *or*, and a scalar implicature triggered by the existential expression *some*. Two groups of 4- to 5-year-olds were recruited for the experiment. Each of these groups was presented with a different implicature type in a between-subjects design. The task of children was to judge whether the test sentences produced by a puppet was correct or incorrect as a description of the final scene in a series of scenes.

The access to alternatives account predicts that children should compute exhaustivity implicatures at a higher rate than they compute scalar implicatures. In contrast, other hypotheses that posit more generalized pragmatic deficits in children – such as a limitation in children's epistemic reasoning or children's pragmatic tolerance – do not predict such a difference in the pattern of children's responses to these different types of inferences. If children do not adhere to the maxim of quantity to the same extent as adults, then this should have an equal effect on both exhaustivity implicatures and scalar implicatures.

Methods

Participants

Twenty-six 4- to 5-year-old children participated in Experiment 1. These children ranged in age from 4;2–5;10 (years;months), with a mean age of 4;4. Children were recruited from a daycare center in Sydney, Australia. We planned to obtain a sample size of 24 child participants, with 12 in each condition as in Barner et al. (2011). In the end, we tested 26 children. Twelve children were tested on a scalar implicature using the existential expression *some*, and 14 were tested with a disjunctive exhaustivity implicature. All children were native speakers of Australian English. Each child was tested individually in a test session that lasted less than five minutes.

Materials

We created eight sets of pictures, four for each between-subject condition. Each set of pictures depicted three different animals, and each condition included four trials. In the exhaustivity condition, two of these animals on each trial had some item or property that the third animal did not have. For example, on one trial, children were shown a tiger, a pig, and a penguin, two of whom had a ball (see Figure 1 below). Two of these characters were mentioned in the disjunctive phrase, including one that did not have the object or property (e.g., *I think the tiger or the pig will have a ball*). In the *some/all* scalar implicature condition, on each trial all three animals had a ball, and children heard an utterance containing *some* (e.g., *I think some of the animals will have a ball*). Table 1 presents an overview of the test sentences and corresponding scene for each condition. Each condition included 4 trials.

Table 1: Overview of test sentences and outcome scene (animal and corresponding object/no object) across experimental conditions

Condition/test sentence	**Situation (outcome scene)**
***some* (scalar implicature)**	
I think some of the animals will have a ball	tiger – has ball
	pig – has ball
	penguin – has ball
***or* (exhaustivity implicature)**	
I think the tiger or the penguin will have a ball	tiger – has ball
	pig – has ball
	penguin – has no object

Procedure

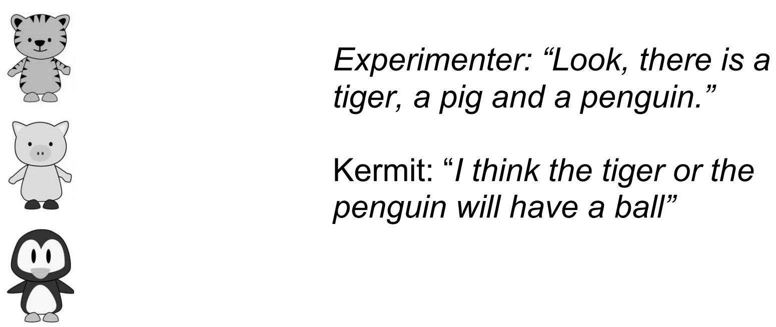
We used a version of the truth value judgment task involving a prediction paradigm (Crain & McKee, 1985; Crain & Thornton, 1998). The two conditions (scalar implicature &

exhaustivity implicature) were presented using a between-subject design, and test sentences were distributed over different items (i.e., in a between-item design). Hence, no item was repeated within participants.

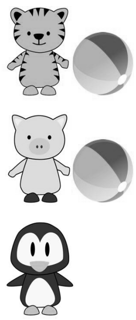
In a warm-up session, children were familiarized with the experimental procedure. The practice trials tested whether children paid attention to the object that appeared on the screen; the first practice trial presented a correct statement concerning the object and a second practice trial an incorrect one. All instructions and test items are provided in Appendix A.1. Items were presented on PowerPoint presentations.

During the test trials, children were presented with statements as predictions about the outcome of a series of scenes. In the first of the series, the experimenter introduced three different animals on a computer screen, for example a tiger, a pig, and a penguin. For the trial type corresponding to Figure 1, the puppet then made a prediction such as “I think the tiger or the penguin will have a ball”. Following the puppet’s prediction, the outcome was revealed (the tiger and the pig were depicted as having a ball), and the child was asked, “*Was Kermit right or wrong?”*. Critically, the scalar alternative (tiger or penguin but not both) was satisfied but the sentence was underinformative because the pig also had a ball. Figure 1 shows a sample item for an exhaustivity implicature and Figure 2 shows an example item for a scalar implicature. Table 2 sketches the structure of one experimental trial. Objects were animated in the PowerPoint presentation.

Scene 1 (character introduction and critical statement)



Scene 2 (outcome)



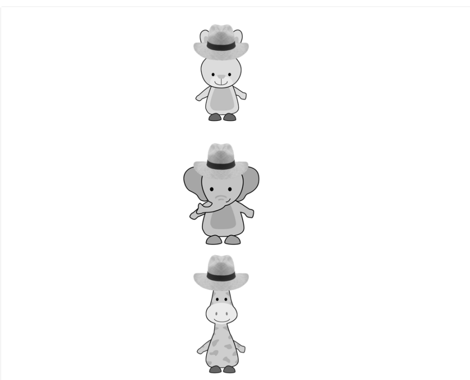
Experimenter: “Was Kermit right or wrong?”

Figure 2: Sample item for exhaustivity implicature, prediction mode scenario

Scene 1 (character introduction and critical statement)

Experimenter: “Look, there is a rabbit, an elephant and a giraffe.”
Kermit: “*I think some of the animals will be wearing a hat”*

Scene 2 (outcome)



Experimenter: “Was Kermit right or wrong?”

Figure 3: Sample item for scalar implicature, prediction mode scenario

Table 2: Structure of one experimental trial

1. Experimenter introduces characters: “Look, there is a donkey, a frog and a cat.”
2. Experimenter introduces prediction: “Let’s see what’s going to happen.”
3. Critical sentence (Kermit’s prediction): “I think only some of the animals will have an apple.”
4. Experimenter clicks to next slide.
5. Experimenter asks for judgment: “Was Kermit right or wrong?”

Results

Prior to conducting the child study, we first conducted a pilot experiment in adult participants to confirm our intuitions regarding exhaustivity implicatures with disjunction, since no previous studies have assessed these in adults. In this pilot, conducted on Amazon Mechanical Turk, participants’ rate of exhaustivity implicatures in *or* statements was 55% (while they computed the scalar implicature associated with *some* at a rate of 80%). Note that because this pilot study differed slightly in design from the child study in order to facilitate online data collection (i.e., we used a truth value judgment task in description mode, rather than the prediction mode used for children that we describe below), we do not report statistical tests comparing adult and child participants. Details on the methodology and results of the adult study are found on the open science framework (https://osf.io/zkqh6/).

Table 4 summarizes the mean percentages of rejections across conditions in Experiment 1. Our primary analyses compared the frequency with which children made judgments that were consistent with *some/all* implicatures (the *some* condition) and exhaustivity implicatures (the *or* condition; see Table 4). Analyses were conducted using Mann Whitney U tests. We found that children made predictions consistent with the computation of implicatures more frequently in the *or* condition (50%) than in the *some* condition (8.3%; $W = 123.5$, $p < .05$). Table B.1 in Appendix B shows the individual rejection rates across conditions per participant. In the *or* condition, children either consistently computed the exhaustivity implicature or never did (except for one child who was inconsistent). In the *some* condition, only one child computed the scalar implicature while all other children did not.

Table 4: Mean % of correct rejections across quantifier condition (Experiment 1)

Quantifier	No *only*
Some (scalar implicature)	8.3 %
Or (disjunctive exhaustivity implicature)	50 %

To interpret children's behavior on the exhaustivity trials, we also examined their justifications for their rejections. This secondary analysis was critical, since a 50% performance indicates that different children responded differently, but leaves open that possibility that this was due to confusion, which lead to guessing rather than to meaningfully different semantic or pragmatic interpretations. Data from children's justifications mitigate this concern, and indicate that children almost always justified their responses by appealing to facts that were in conflict with the inferences of the test sentences. Further, it is crucial to note that individual children were consistent in their responses. In other words, children did not merely respond randomly to the test sentences.

On the 50% of exhaustivity trials on which children rejected the puppet's underinformative statement (N = 28 trials), 62% of their justifications pointed out that both the tiger and the pig had a ball and a further 17% of children's justifications simply pointed out that the pig had a ball. These justifications are compatible with children having computed an implicature – i.e., that the penguin or a tiger but no other animal should have a ball – though alternative analyses, which we return to in the General Discussion, may also be possible. Critically, few of children's justifications (7%) mentioned the fact that the penguin did not have a ball. This finding fails to support the hypothesis that children license a conjunctive inference in response to sentences with disjunction (Singh et al., 2016; Tieu et al., 2016). If disjunction licensed a conjunctive inference for children, then the fact that both the pig had a ball and the penguin did *not* have a ball should have led children to reject the test sentences with disjunction. However, children were substantially more likely to mention the pig. This finding suggests that, if children compute conjunctive inferences at all, this tendency is very weak. Another 10% of children's justifications for rejecting the puppet's statements pointed out that the tiger and the pig had a ball but the penguin did not have one. Because this is merely a redescription of the entire scene, it does not favor any particular interpretation over another – a child who only computes an exhaustivity inference might redescribe the scene to explain their judgment, as might a child, who licenses a conjunctive inference. Finally, 3.5% of children's justifications were unintelligible.

In the *some* condition, only one child rejected the puppet's underinformative statements. This child pointed out on all four trials that all of the animals had the property mentioned in the puppet's statement, so this child's justifications were clearly consistent with the successful computation of a scalar implicature. Thus, the percentages of children's judgments across the conditions, and children's justifications for rejecting the puppets statements, both demonstrated a strong difference in their capacity to generate the two kinds of implicatures.

Discussion

Experiment 1 compared 4- to 5-year-olds' performance on scalar implicatures (some but not all) and exhaustivity implicatures (*a* or *b* but nobody else) in a prediction mode paradigm of the truth value judgment task. The finding that children experienced significant difficulty computing *some/all* implicatures replicated a now-standard finding in the literature. In contrast, we found that children preformed significantly better in computing exhaustivity implicatures with disjunction. The comparison across conditions is in keeping with the access to alternatives account. According to this account, 4- to 5-year-olds have acquired the mechanism of exhaustification but have difficulties spontaneously accessing alternatives that are not made salient in the context, or are substrings of the test sentences. Consistent with our assumption that children computed an exhaustivity implicature, their justifications showed that they rejected the *or* statements on the grounds that the unmentioned third character had the same property as one of the disjuncts that was mentioned in the test sentence. Therefore, the hypothesis which is most in line with the findings from Experiment 1 is that children struggled to automatically access alternatives from the mental lexicon, but they were able to access explicitly mentioned material (cf. Barner et al., 2011).

Experiment 2

The main hypothesis of Experiment 1 was that if access to alternatives constraints children's ability to compute implicatures, then they should be more likely to compute *or* exhaustivity inferences than *some/all* implicatures. Consistent with this, we found that whereas children computed exhaustivity inferences 50% of the time, they almost never computed scalar implicatures involving the lexical *some/all* scale. While this result is consistent with our hypothesis, it leaves open several important questions. In particular, there are two potential issues with Experiment 1, which we will address in the following experiment.

First, in Experiment 1 we used statements with embedding under *I think*, which may have lowered rates of implicature.[4] Although there are *a priori* reasons to believe that children may interpret such statement similarly to unembedded statements in experiments that use prediction mode, there is not empirical data to confirm this. Further, whereas Experiment 1 used the *some/all* scale as a control, a more direct comparison is possible using the *or/and* scale. Experiment 2 therefore sought to replicate the findings of Experiment 1 with two modifications: (i) we tested children with unembedded statements (e.g., *The tiger or the penguin will have a ball)* and (ii) we tested the scalar implicature triggered by disjunction as a control case (*The tiger or the penguin but not both*).

A second goal of Experiment 2 was to investigate exhaustivity implicatures in conjunctive statements (e.g., *The tiger and the penguin will have a ball*). In the disjunctive case, the subdomain alternatives are more informative than the disjunctive statement itself. Hence, participants may reason why the puppet has not uttered a more informative simpler statement

[4] Work by Geurts & Pouscoulous (2009) has called the existence of embedded implicatures into question. Yet there is now ample experimental evidence for the existence of implicatures under different kinds of embedding operators (e.g. Chemla & Spector, 20011; Potts et al., 2016; Gotzner & Benz, 2018: Gotzner & Romoli, 2018).

about a single individual. So, participants may consider each individual disjunct as well as the additional contextual alternative (*The pig will have a ball*). In conjunctive statements, on the other hand, the subdomain alternatives are not stronger than the original statement, and as a consequence they should not be used in inference computation.[5] Participants may thus focus on whether the two mentioned individuals have the property, which is verified by the critical outcome scene in our experiment. For this reason, we predict that children should be less likely to compute exhaustivity implicatures with conjunction compared to disjunction.

Methods

Participants

Forty-eight 4- to 5-year-olds were recruited at a daycare center in Sydney, Australia. The child participants ranged in age from 4;0 to 5;11, with a mean age of 4;7. All children were native speakers of Australian English. In addition, 26 children were recruited at a daycare center in San Diego, ranging in age from 4;0 to 5;10, with a mean age of 4;7.[6] All children tested in San Diego were native speakers of American English. Each child was tested individually, and a test session lasted about five minutes.

Materials

Experiment 2 used the picture materials from Experiment 1 and compared three sentence types: *or* (disjunctive exhaustivity implicature), *and* (conjunctive exhaustivity implicature) and *or* (scalar implicature) (see Table 5). All conditions were between subjects. Participants in the experiment were presented with four trials per condition interspersed with four filler items (patently false or patently true), totaling in eight experimental trials.

---

[5] A further difference between disjunctive and conjunctive statements is that the *ad hoc* scalar alternative that participants might construct would be more complex than the original statement in the conjunctive case (e.g., *The tiger and the penguin and the pig will have a ball*). In the disjunctive case, on the other hand, the original statement and potential *ad hoc* alternatives are of equal complexity (e.g., *The tiger and pig will have a ball*). Following the structural account of alternatives by Fox and Katzir (2011), alternatives should be at most as complex as the prejacent. Therefore, the conjunctive alternative involving all three individuals may not be derived.

[6] The scalar implicature control condition with disjunction was carried out in San Diego. Prior to testing these children, we first conducted a partial replication of the exhaustivity condition with disjunction, and found rejection rates (42%) comparable to those found in Sydney (50%).

In a warm-up phase, children were familiarized with the experimental procedure. The first practice trial tested mention of the object that appeared on the screen and a second practice trial presented an incorrect situation with a different object. All instructions and test items are found in Appendix A.2. We prepared PowerPoint presentations with one order of presentation of the stimuli.

Table 5: Overview of test sentences and outcome scene (animal and corresponding object) across experimental conditions (Experiment 2)

Condition/test sentence	Situation (outcome scene)
***or* (disjunctive exhaustivity implicature)**	
The tiger or the penguin will have a ball	tiger – has ball
	pig – has ball
	penguin – has no object
***and* (conjunctive exhaustivity implicature)**	
The tiger and the penguin will have a ball	tiger – has ball
	pig – has ball
	penguin – has ball
***or* (scalar implicature)**	
The tiger or the penguin will have a ball	tiger – has ball
	pig – has no object
	penguin – ball

Procedure

The procedure was the same as in Experiment 1.

Results

Prior to running Experiment 2, we conducted a second online pilot study, comparing exhaustivity implicature with conjunction and disjunction. In this study, adults rejected disjunctive statements at a rate of 45% while only 5% of participants rejected conjunctive statements in a corresponding non-exhaustive situation. Again, due to differences in methodology, we do not compare the results statistically. Details on the second pilot study are found on https://osf.io/ahs45/.

Table 7 summarizes the mean percentages of rejections across conditions for Experiment 2. One item (out of four) in the *or* and *and* exhaustivity conditions showed the wrong object and was therefore excluded from further analyses. Patently true filler trials were correctly accepted in 97.4% of the cases and patently false sentences were correctly rejected 91.9%, confirming that kids understood the task.

Table 7: Mean % of rejections across conditions (Experiment 3)

Condition	% rejected
Or (disjunctive exhaustivity implicature)	55 %
And (conjunctive exhaustivity implicature)	22 %
Or (scalar implicature)	14 %

We compared mean ratings across conditions with a series of Mann Whitney U tests. To account for multiple comparisons, we performed a Bonferroni correction, resulting in an adjusted alpha level of .017. These analyses found that participants computed significantly more exhaustivity implicatures with disjunction compared to conjunction (*or* exhaustivity 55% vs. *and* exhaustivity 22%; W = 408, $p < .01$) and compared to the scalar implicature of disjunction (*or* scalar 12 %; W = 480.5, $p = .001$). The difference between conjunction and the scalar implicature, in turn, was not significant (W = 268, $p = 0.273$).

Again, a substantial number of children rejected the underinformative *or* statement with rates comparable to that of Experiment 1 (Exp.1: 50%, Exp.2: 55%). Participants were more likely to compute exhaustivity with disjunction compared to conjunction and they were least likely to derive the scalar implicature of disjunction. Table C.1 in Appendix B shows the individual rejection rates across conditions per participant.

We again analyzed children's justifications of the trials in which they rejected the underinformative statements (*or* exhaustivity: 40 trials; *and* exhaustivity: 16 trials; *or* scalar: 15 trials). In the *or* exhaustivity condition, 50% of the children pointed out that the tiger and the pig had a ball, 20% said that the pig had a ball too, 12.5% pointed out that the tiger and the pig had a ball while the penguin did not have one and 17.5% of the responses were unintelligible. These justifications again indicate that the majority of child participants rejected the underinformative disjunctive statements because the additional mentioned character had the property one of the disjuncts had. This is evidence that children computed an exhaustivity implicature from the disjunctive statements. On the other hand, none of the children pointed out that the penguin did not have a ball, again contrary to the hypothesis that children interpret *or* conjunctively.

In the *and* exhaustivity condition, 50% of the children (who rejected the statement) pointed out that all animals had a ball, 25% said that the pig had a ball too and 25% of the responses were unintelligible. These justifications indicate that some children computed an exhaustivity implicature in the conjunctive statement (though on fewer trials than in the disjunctive condition).

In the *or* scalar condition, 7% pointed out that the tiger and penguin had a ball (again failing to find evidence for conjunctive reading), 27% of the rejections were justified by saying that only two animals had a ball, 27% said that the pig did not have a ball, 7% pointed out that not all animals have a ball and another 33% of the responses were unintelligible.

Overall, the results show that children computed an exhaustivity implicature in disjunctive statements while they were less likely to compute this inference in conjunctive statements. Further, kids were also less likely to compute the scalar implicature of disjunction.

Discussion

In Experiment 2, we used unembedded statements with disjunction and conjunction and to test whether children derive exhaustivity implicatures for either type of statement and whether they derive a scalar implicature for disjunctive statements. First, we replicated the finding that children are more likely to derive exhaustivity implicatures than scalar implicatures for the same lexical trigger, that is disjunction. Further, we found that children are more likely to derive exhaustivity implicatures with disjunction than with conjunction, with adults behaving similarly in this respect.

We assume that the difference in the availability of exhaustivity implicatures for disjunctive and conjunctive statements follows from considerations of informativeness in the original statements and their alternatives. When the puppet uses a disjunctive statement, he does not predict which of the mentioned individuals will have a ball. However, participants can use the puppet's statement to infer that no additional character in the display should have one. Alternative statements about the individual characters would have been informative (and also more simple). If, on the other hand, the puppet utters a conjunctive statement, he makes a correct prediction about the outcome scene, giving participants a reason to accept the statement. In conjunctive statements, the subdomain alternatives are not stronger than the original statement and neither is the alternative statement about the additionally-mentioned character. For this reason, exhaustivity implicatures are less available with conjunction than disjunction. We turn to a discussion of the relevance of these findings for different accounts of implicature in the following sections.

Conjunctive sentences like *The tiger and the penguin will have a ball* also have S(A) and S(B) as subdomain alternatives but these are not informationally stronger than the original statement and neither is the additional contextual alternative S(C). That is, the only informationally stronger alternative would be an *ad hoc* scale involving all three alternatives S(A and B and C) but this is of higher complexity. Under the assumption of the structural account of alternatives by Fox and Katzir (2011) alternatives should be at most as complex as the prejacent, hence the alternative S(A and B and C) will not be derived. Participants should therefore focus on checking whether the two mentioned individuals have the property when presented with conjunctive statements. Since this situation is verified in the outcome scene, participants should accept the statement.

General Discussion

We conducted two experiments evaluating children's ability to compute three forms of implicature. First, Experiment 1 tested exhaustivity implicature triggered by disjunctive statements, and scalar implicature triggered by the quantifier *some*. The main finding was that children performed better on exhaustivity implicatures compared to the *some/all* implicatures. More specifically, while 50% of the children consistently computed exhaustivity implicatures, only a single child provided evidence of having computed a *some/all* implicature. We reasoned that these results were likely due to the fact that children have difficulties accessing relevant alternatives for lexicalized scales. Second, Experiment 2 compared three trial types – exhaustivity implicatures in disjunctive vs. conjunctive statements and scalar implicature with the *or/and* scale. We again found that children were more likely to compute the exhaustivity implicature with *or* than the corresponding scalar implicature. Interestingly, we also found that statements with *and* triggered an exhaustivity implicature to a lesser extent (with a similar pattern for adult speakers).

The results of these two experiments provide new evidence that, whereas children have acquired the mechanism necessary to compute implicatures, they struggle to do so whenever they are required to access scalar alternatives which are not presented in the context (Barner et al., 2011). These findings go beyond the findings of Barner et al. (2011) and subsequent studies in three ways. To begin, whereas several studies have shown that children can compute various inferences that resemble implicatures in forced choice paradigms (Miller et al., 2005; Stiller et al., 2015; see also Papafragou & Tantalou, 2004, for a study in which alternatives are explicitly provided), few have asked whether children can spontaneously construct *ad hoc* alternatives in order to accept or reject statements with regard to a particular state of the world. Given that this is generally the scenario in which children struggle to compute implicatures (e.g., Noveck, 2001; Papafragou & Musolino, 2003), it is critical to know which factors affect performance under related conditions. Likewise, whereas Barner et al. (2011) showed unequivocally that access to alternatives affects children's behaviors, they did so in a task that asked children questions, and thus did not trigger implicatures (e.g., the question *Are the dog and the cat sleeping?* was answered with a "yes" response by children and adults).

The present study filled this gap by demonstrating that children are capable of deriving exhaustivity implicatures from descriptive statements. Also, we looked at a novel test case: Disjunction presented in a context where a third character had the same property as one of the disjuncts. We found that children rejected the statements in this situation on the grounds that the other character had a relevant property too. Therefore, we concluded that children are able to consider alternatives from the context and use these to compute the corresponding inferences of an utterance. Further, we showed that the type of statement affects whether children consider these additional contextual alternatives. In disjunctive statements, children used the statement to infer that the additional mentioned character does not have the relevant property while not inferring the scalar implicature that *not both* have it. In the case of conjunctive statement, on the other hand, the statement was informative with respect to the two mentioned characters having a ball and thus children were less likely to compute an exhaustivity implicature. These results indicate that children do not build an *ad hoc* scalar alternative that is more complex than the original statement (for conjunctive statements: S(A and B and C)). Rather, exhaustivity implicatures are most likely to occur via negation of a simple statement about a contextually-mentioned character.

Ramifications for different accounts

Overall, our results support the idea that children's ability to compute implicatures is constrained by their ability to access relevant alternative utterances. Therefore, the results support the access to alternatives account, over accounts which appeal to limits in theory of mind reasoning, working memory, or pragmatic tolerance. The results are also consistent with the view that the same mechanism of exhaustification underlies superficially distinct phenomena such as scalar implicatures, exhaustivity implicatures, and free choice inferences, and suggest that children have acquired this mechanism of exhaustification, beginning at least by the age of 4 (see also Tieu et al., 2016).

As we noted earlier, Singh et al. (2016) account for differences in the interpretations assigned to disjunctive statements by children and adults by proposing that children, but not adults, license a conjunctive inference in comprehending disjunctive statements (for a similar claim, see Tieu et al., 2015). These conjunctive inferences are attributed to children's difficulty in accessing the scalar alternative *and*. Lacking access to this alternative, children's alternatives consist of only the subdomain alternatives, namely statements that mention just one of the disjuncts. Our study, however, did not find evidence of conjunctive inferences among children. While it was not specifically designed to test the conjunctive pattern, it would be expected that children in our exhaustive *or* condition should reject statements because only one of the disjuncts is true. Although children often exhibited inclusive readings of disjunction (e.g., accepting a disjunctive statement when both disjuncts were true), only 7% of rejections were justified by pointing out that one of the mentioned characters (e.g., the penguin) did not have the relevant property. This is far fewer than the level of approximately 50% reported by both Singh et al. and Tieu et al. (for similar failures to find evidence that children computed conjunctive inferences in interpreting sentences with disjunction, and a potential explanation, see Huang & Crain, forthcoming; Skordos, Feiman, Bale, & Barner, under review). Future studies should explore the nature of this difference.

We should also highlight the fact that our findings do not provide evidence against a Gricean account of implicature derivation *per se*. Rather, they provide evidence against the explanation that children fail to derive implicatures because they tolerate violations of the maxim of quantity. Proponents of the Gricean account could potentially invoke a similar explanation that children require relevant alternatives to be contextually salient. While this study does not directly test the role of Gricean reasoning, it does show, however, that a significant proportion of children's difficulty can be explained by other factors. Also, previous studies show that children can compute ignorance implicatures (e.g., Hochstein et al., 2016), which involve similar Gricean abilities invoked for the computation of scalar implicatures. Still, the ability to infer alternatives are informative can be characterized as a component of Gricean reasoning. The current results show that children do not lack these abilities entirely. However, children may require the speaker's intent to be made more explicit and consistent within a single situation (see also Skordos & Papafragou, 2014).

As discussed in the introduction, many accounts of scalar inference propose that scalar implicature in the case of disjunction is due to competition between a statement including *or* and a more informative alternative statement containing *and*. We showed that children do not readily access this alternative. There are alternative accounts that derive scalar inferences through the exhaustification of subdomain alternatives (e.g., van Rooij & Schulz, 2004; Schulz & van Rooij, 2006; Spector, 2006). Such accounts would derive both the scalar inference as well as the exhaustivity inference of disjunction through the same mechanism. Our study provided evidence that children have the capacity to negate contextually-mentioned alternatives and that they are guided by considerations of informativeness. Under this account, children would arrive at the exhaustivity implicature by negation of the subdomain alternative S(C) about the contextually-mentioned character. This is likely to occur for disjunctive statements because the original statement is not informative regarding which of the two mentioned individuals will have the relevant property.

We have also looked at a novel test case that has not been investigated to date. Conjunctive sentences like *The tiger and the penguin will have a ball* also have S(A) and S(B) as subdomain alternatives but crucially these are not informationally stronger than the original statement (and neither is the additional contextual alternative S(C)). The only informationally stronger alternative would be an *ad hoc* scale involving all three alternatives S(A and B and C) but this is of higher complexity. Our results provide evidence that such an *adhoc* scale is not readily derived, in line with the view by Fox and Katzir (2011) that alternatives are at most as complex as the prejacent.

As a final point, we note that our data, like many in the literature on implicature, do not show direct evidence of scalar implicatures *per se*, since children might conceivably have made judgements basis on the underinformativity of the weaker interpretation, rather than the fact that a weak statement implies the negation of statements with stronger alternatives (for discussion, see Katsos & Bishop, 2016). For example, when presented with *The tiger or the penguin will have a ball*, children might reject this utterance because a more informative utterance might have been uttered (i.e., that mentions the pig), but not because it implies the negation of that stronger alternative. While we acknowledge this possibility, it remains the case that if access to alternatives affects children's ability to compute inferences on the basis of informativity, it must also affect their ability to compute any inference that relies on informativity, such as implicature. Thus, studies in this literature contribute to the question of why children struggle to compute implicatures, even if certain measures are ambiguous with respect to whether or not scalar alternatives are negated during interpretation.

Conclusions

We conducted two experiments using a truth value judgment task to compare 4- to 5- year-old children's performance on exhaustivity implicatures and scalar implicatures. The findings of these experiments demonstrated that 4- to 5- year-old children are able to compute exhaustivity implicatures with disjunction, whereas they successfully compute scalar implicatures to a lesser extent. Further, we showed that children are more likely to compute exhaustivity implicatures with disjunctive compared to conjunctive statements. Taken together, the findings of these experiments support the conclusion that children are able to compute implicatures, but do not automatically access scale mates. We interpret the findings as evidence supporting the access to alternatives account, rather than more general pragmatic deficits.

Acknowledgments

We would like to express our gratitude to Cory Bill, Andreas Haida, Paul Marty, Clemens Mayr, Andreea Nicolae, Uli Sauerland, Stephanie Solt, Benjamin Spector, Bob van Tiel, Jack Tomlinson and Judith Tonhauser for helpful discussion. We also thank Elena Pagliarini and Cory Bill for providing us with the picture stimuli. We are very grateful to Elena D'Onofrio and Ashlie Pankonin for assistance with data collection.

References

Antoniou, K., Cummins, C., & Katsos, N. (2016). Why only some adults reject under-informative utterances. *Journal of Pragmatics, 99*, 78-95.

Barner, D., Chow, K., & Yang, S. J. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology*, *58*(2), 195-219.

Barner, D. & Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology*, *60*, 40-62.

Barner, D., Brooks, N. & Bale, A. (2011). Accessing the unsaid: the role of scalar alternatives in children's pragmatic inference. *Cognition*, *118*, 84-93.

Barner, D., Hochstein, L., Rubenson, M., and Bale, A. (2018). Four-year-old children compute scalar implicatures in absence of epistemic reasoning. In Syrett, K.; and Arunachalam, A., editor(s), Semantics in Acquisition. Amsterdam: Benjamins. pp. 326-349.

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

Braine, M., & Rumain, B. (1981). Children's comprehension of or: evidence for a sequence of competencies. *Journal of Experimental Child Psychology, 31*, 46-70.

Chierchia, G. (2006). Broaden your Views. Implicatures of Domain Widening and the "Logicality" of Language. *Linguistic Inquiry*, *37*(4), 535-590.

Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A. & Meroni, L. (2001). The acquisition of disjunction: evidence for a grammatical view of scalar implicatures. In: A. H. J. Do, L. Dominguez & A. Johanson (Eds.), *Proceedings of the 25th Boston University Conference on Language Development*. Cascadilla Press: Somerville, MA, 157-168.

Chierchia, Fox, Spector (2012). Scalar implicature as a grammatical phenomenon. In: *Semantics: An International Handbook of Natural Language Meaning*. Vol. 3. Edited by Maienborn, von Heusinger & Portner. Mouton de Gruyter, 2297-2331.

Crain, S. & McKee, C. (1985): The acquisition of structural restriction on anaphora. In: S. Bresnan, J. Choe & J. McDonough (Eds.), *Proceedings of the 15th Annual Meeting of the North Eastern Linguistic Society.* Graduate Linguistic Student Association: Amherst, MA, 94-110.

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

De Neys, Wim & Walter Shaeken. 2007. When people are more logical under cognitive load: dual task impact on scalar implicature. *Experimental Psychology 54(2*). 128–133.

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect.*Merrill-Palmer Quarterly*, *50*(3), 274-290.

Foppolo, F., Guasti, M. T., & 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: U. Sauerland & P. Stateva (Eds.), *Presupposition and implicature in compositional semantics*. Palgrave Macmillan: New York, 537-586.

Fox, D., & Katzir, R. (2011). On the characterization of alternatives. *Natural Language Semantics*, *19*(1), 87–107.

Gazdar, G. (1979). Pragmatics: Implicature, presupposition, and logical form. New York: Academic Press.

Gotzner, N., & Benz, A. (2018). The best response paradigm and a comparison of different models of implicatures of complex sentences. *Frontiers in Communication, 2 (21)*.

Gotzner, N. & Romoli, J. (2018). The scalar inferences of strong scalar terms under negative quantifiers and constraints on the theory of alternatives. *Journal of Semantics, 35 (1)*, 95-126.

Grice, P. (1975). Logic and conversation. *Syntax and Semantics*, *3(41),* 58.

Gualmini, A., Crain, S., Meroni, L., Chierchia, G., & Guasti, M. T. (2001, October). At the semantics/pragmatics interface in child language. In *Semantics and Linguistic Theory*, *11*, 231-247.

Guasti, M., Chierchia, G., Crain, S., Foppolo, F., Gualmini, A., & Meroni, L. (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes*, *20*(5), 667–696.

Geurts, Bart & Nausicaa Pouscoulous. 2009a. Embedded implicatures?!? *Semantics & Pragmatics 2(4)*. 1–34. doi:10.3765/sp.2.4.

Hochstein, L., Bale, A. & Fox, 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.

Horowitz, A. C., Frank, M. C. (2014). Preschoolers infer contrast from adjectives if they can access lexical alternatives. *Proceedings of the 36th Annual Conference of the Cognitive Science Society*, 625-631, Quebec.

Huang, Y. T., & Snedeker, J. (2009). Semantic meaning and pragmatic interpretation in 5-year-olds: evidence from real-time spoken language comprehension. *Developmental psychology*, *45*(6), 1723.

Huang, Y. T., & Snedeker, J. (2009b). Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive Psychology*, *58*(3), 376-415.

Katsos, N. & Bishop, D. V. M. (2011). Pragmatic Tolerance: Implications for the Acquisition of Informativeness and Implicature. *Cognition*, *20*, 67-81.

Levinson, S. C. (2000): Presumptive meanings: The theory of generalized conversational implicature. MIT press: Cambridge.

Miller, K., Schmitt, C., Chang, H., & Munn, A. (2005). Young children understand some implicatures. In: *Proceedings of the 29th Annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press, 389–400.

Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*(2), 165–188.

Noveck, I. A., & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language*, *85*(2), 203-210.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: experiments at the semantics-pragmatics interface. *Cognition, 86*, 253–282.

Papafragou, A., & Tantalou, N. (2004). Children's computation of implicatures. *Language Acquisition, 12*, 71–82.

Paris, S. (1973). Comprehension of language connectives and propositional logical relationships. *Journal of Experimental Child Psychology, 16*, 278-291.

Pouscoulous, N., Noveck, I. A., Politzer, G., & Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition*, *14*(4), 347–375.

Potts, C., Lassiter, L., Levy, R. & Frank. M.C. (2016) Embedded implicatures as pragmatic inferences under compositional lexical uncertainty. *Journal of Semantics 33(4)*, 755-802.

Reinhart, T. (2004). The processing cost of reference-set computation: Acquisition of stress shift and focus. *Language Acquisition, 12*(2), 109–155.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy*, 27 (3), 367-391.

Skordos, D. & Papafragou, A. (2014). Scalar Inferences in 5-year-olds: The Role of Alternatives. *Proceedings from the 38th Annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press.

Skordos, D., Feiman, R., Bale, A. & Barner, D. (under review). Do children interpret *or* conjunctively? https://psiarxiv.com/4uqn9/.

Smith, C. L. (1980). Quantifiers and question answering in young children. *Journal of Experimental Child Psychology, 30*, 191-205.

Stiller, A. J., Goodman, N. D., & Frank, M. C. (2015). Ad-hoc implicature in preschool children. *Language Learning and Development, 11*(2), 176-190.

Singh, R., Wexler, K., Astle, A., Kamawar, D., & Fox, D. (2016). Children interpret disjunction as conjunction: Consequences for the theories of implicature and child development. Accepted with revisions, *Natural Language Semantics,* 24(4).

Tieu, L., Romoli, J., Zhou, P., & Crain, S. (2015). Children's Knowledge of Free Choice Inferences and Scalar Implicatures. *Journal of Semantics*, doi: 10.1093/jos/ffv001.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*, 749-750.

Appendix A.1: Instructions for Experiment 1

Let's take a look at the following pictures and see what Kermit thinks about them. We will ask Kermit what's going to happen and he tells us what he thinks. Then, you tell me whether he was right or wrong.

Training (practice trials)

1. Experimenter: Look, there are bears. Let's see what's going to happen.

   Kermit: I think the bears will have cookies.

   (Click to next slide).

   Was Kermit right?

2. Experimenter: Look, there are pigs. Let's see what's going to happen.

   Kermit: I think the pigs will have cars.

   (Click to next slide).

   Was Kermit right?

Test (critical trials)

3. Experimenter: Look, there is a tiger, a pig and a penguin. Let's see what's going to happen.

   Kermit: I think the tiger or the penguin will have a ball.

   (Click to next slide).

   So, was Kermit right?

   Was Kermit right?

4. Experimenter: Look, there is a tiger, a rabbit and a penguin. Let's see what's going to happen.

   Kermit: I think the tiger or the penguin will have a milkshake.

   (Click to next slide).

   Was Kermit right?

5. Experimenter: Look, there is a hippo, a deer and a chicken. Let's see what's going to happen.

Kermit: I think the hippo or the chicken will have a bucket.

(Click to next slide).

Was Kermit right?

6. Experimenter: Look, there is a lion, a rabbit and a kangaroo. Let's see what's going to happen.

   Kermit: I think the lion or the kangaroo will have a dumbbell.

   (Click to next slide).

   Was Kermit right?

Appendix A.2: Instructions for Experiment 2

Let`s see what Kermit thinks about the following pictures. We will ask Kermit what`s going to happen and he tells us what he thinks. Then, you tell me whether he was right or wrong.

Training

1. Experimenter: Look, there are bears. Let`s see what`s going to happen.

   Kermit: **The bears will have cookies.**

   (Click to next slide).

   Was Kermit right?

2. Experimenter: Look, there are pigs. Let`s see what`s going to happen.

   Kermit: **The pigs will have cars.**

   (Click to next slide).

   Was Kermit right?

Test

3. Experimenter: Look, there is a tiger, a pig and a penguin. Let`s see what`s going to happen.

   Kermit: **The tiger or the pig will have a ball.**

   (Click to next slide).

   Was Kermit right?

4. Experimenter: Look, there is a bear, a chicken and an elephant. Let`s see what`s going to happen.

   Kermit: **The animals will have a bike.**

   (Click to next slide).

   Was Kermit right?

5. Experimenter: Look, there is a lion, a rabbit and a kangaroo. Let`s see what`s going to happen.

   Kermit: **The lion or the rabbit will have a book.**

   (Click to next slide).

Was Kermit right?

6. Experimenter: Look, there is a rabbit, an elephant and a giraffe. Let`s see what`s going to happen.

   Kermit: **The animals will be wearing a hat.**

   (Click to next slide).

   Was Kermit right?

7. Experimenter: Look, there is a hippo, a deer and a chicken. Let`s see what`s going to happen.

   Kermit: **The hippo or the deer will have bucket.**

   (Click to next slide).

   Was Kermit right?

8. Experimenter: Look, there is a lion, a pig and a koala. Let`s see what`s going to happen.

   Kermit: **The animals will catch a butterfly.**

   (Click to next slide).

   Was Kermit right?

9. Experimenter: Look, there is a tiger, a rabbit and a penguin. Let`s see what`s going to happen.

   Kermit: **The tiger or the rabbit will have a milk shake.**

   (Click to next slide).

   Was Kermit right?

10. Experimenter: Look, there is a mouse, a zebra and a monkey. Let`s see what`s going to happen.

    Kermit: **The zebra and the monkey will have a candy.**

    (Click to next slide).

    Was Kermit right?

Appendix B

Table B.1: Individual participant means across conditions (Experiment 1)

participant	condition	mean
1	*some*	0.00
2	*some*	1.00
3	*some*	0.00
4	*some*	0.00
5	*some*	0.00
6	*some*	0.00
7	*some*	0.00
8	*some*	0.00
9	*some*	0.00
10	*some*	0.00
11	*some*	0.00
12	*some*	0.00
13	*or*	0.00
14	*or*	1.00
15	*or*	1.00
16	*or*	0.75
17	*or*	0.50
18	*or*	0.00
19	*or*	1.00
20	*or*	0.00
21	*or*	0.00
22	*or*	1.00
23	*or*	0.00

24	*or*	0.00
25	*or*	1.00
26	*or*	0.75

Table C.2: Individual participant means across conditions (Experiment 3)

participant	condition	mean
1	or exhaustivity	1
2	or exhaustivity	1
3	or exhaustivity	0.33
4	or exhaustivity	0
5	or exhaustivity	1
6	or exhaustivity	1
7	or exhaustivity	0.33
8	or exhaustivity	0.33
9	or exhaustivity	1
10	or exhaustivity	1
11	or exhaustivity	1
12	or exhaustivity	0
13	or exhaustivity	0
14	or exhaustivity	0
15	or exhaustivity	1
16	or exhaustivity	0
17	or exhaustivity	1
18	or exhaustivity	1
19	or exhaustivity	1
20	or exhaustivity	0
21	or exhaustivity	0.33
22	or exhaustivity	0.67
23	or exhaustivity	0.33
24	or exhaustivity	0
25	and exhaustivity	1
26	and exhaustivity	0
27	and exhaustivity	0
28	and exhaustivity	1
29	and exhaustivity	0
30	and exhaustivity	0.33
31	and exhaustivity	0
32	and exhaustivity	0

33	and exhaustivity	0.33
34	and exhaustivity	1
35	and exhaustivity	0
36	and exhaustivity	1
37	and exhaustivity	0
38	and exhaustivity	0
39	and exhaustivity	0
40	and exhaustivity	0
41	and exhaustivity	0
42	and exhaustivity	0
43	and exhaustivity	0
44	and exhaustivity	0.33
45	and exhaustivity	0
46	and exhaustivity	0.33
47	and exhaustivity	0
48	and exhaustivity	0
49	or scalar	1
50	or scalar	0.75
51	or scalar	0
52	or scalar	0
53	or scalar	0
54	or scalar	0
55	or scalar	0
56	or scalar	0
57	or scalar	0
58	or scalar	0
59	or scalar	0
60	or scalar	0
61	or scalar	0
62	or scalar	0
63	or scalar	0
64	or scalar	0
65	or scalar	0
66	or scalar	0
67	or scalar	0

68	or scalar	0.75
69	or scalar	0.25
70	or scalar	1
71	or scalar	0
72	or scalar	0
73	or scalar	0
74	or scalar	0