Do children interpret "or" conjunctively?

Article *in* Journal of Semantics · November 2019

CITATION
1

READS
78

4 authors:

Dimitrios Skordos
The University of Calgary
12 PUBLICATIONS **65** CITATIONS
SEE PROFILE

Roman Feiman
Brown University
8 PUBLICATIONS **330** CITATIONS
SEE PROFILE

Alan Bale
Concordia University Montreal
46 PUBLICATIONS **513** CITATIONS
SEE PROFILE

David Barner
University of California, San Diego
144 PUBLICATIONS **1,734** CITATIONS
SEE PROFILE

Do children interpret “or” conjunctively?

Dimitrios Skordos[a], Roman Feiman[b], Alan Bale[c], & David Barner[d]

[a]University of Calgary

[b]Brown University

[c]Concordia University

[d]University of California, San Diego

CONTACT:

Dimitrios Skordos

School of Languages, Linguistics, Literatures & Cultures

University of Calgary

dimitrios.skordos@ucalgary.ca

Acknowledgements: Thank you to the many families and preschools who participated in this research, as well as the Fleet Science Center. Thank you also to Junyi Chu and members of the Language and Development Lab for help collecting data. This work was funded in part by a grant to D.B. from the James S. McDonnell Foundation and a SSHRC Insight grant to A.B.

ABSTRACT

Preschoolers often struggle to compute scalar implicatures (SI) involving disjunction (*or*), in which they are required to strengthen an utterance by negating stronger alternatives, e.g., to infer that, “The girl has an apple or an orange” likely means she doesn’t have both. However, recent reports surprisingly find that a substantial subset of children interpret disjunction *as conjunction*, concluding instead that the girl must have both fruits. According to these studies, children arrive at conjunctive readings not because they have a non-adult-like semantics, but because they lack access to the stronger scalar alternative *and,* and employ doubly exhaustified disjuncts when computing implicatures. Using stimuli modeled on previous studies, we test English-speaking preschoolers and replicate the finding that many children interpret *or* conjunctively. However, we speculate that conditions which replicate this finding may be pragmatically infelicitous, such that results do not offer a valid test of children’s semantic competence. We show that when disjunctive statements are uttered in contexts that render the speaker’s intended question more transparent, conjunctive readings disappear almost entirely.

Older preschoolers and kindergarteners often struggle to compute scalar implicatures (SIs) involving quantifiers like *some* (Smith, 1980; Noveck, 2001; Papafragou & Musolino, 2003) and logical operators like disjunction (Braine & Rumain, 1981; Chierchia et al., 2001). For example, when presented with an utterance containing *some* like the one in (1a), many children fail to derive the implicature in (1c). Similarly, when presented with a sentence like the one in (2a) many children fail to derive an exclusive interpretation of *or* characterized by (2c):

(1) a. The boy took some of the bananas.

b. The boy took all of the bananas.

c. The boy took some, but not all of the bananas.

(2) a. The girl has an apple or an orange.

b. The girl has an apple and an orange.

c. The girl has an apple or an orange, but not both.

On most accounts, deriving a scalar implicature involves accessing a stronger alternative statement that is generated by replacing one scalar term (e.g., *some*, *or*) with another, stronger, one (e.g., *all*, *and*). For example, to derive an implicature for the sentence containing *some* in (1a), children must access and negate the stronger alternative in (1b). Likewise, to derive the implicature for the sentence containing *or* in (2a), children must access and negate the stronger alternative in (2b). On standard accounts of disjunction, a failure to do so should result in an inclusive interpretation, wherein the listener concludes that the utterance in (2a) is acceptable even when the girl has both an apple *and* an orange (Crain & Khlentzos, 2010; Gazdar, 1979; McCawly, 1993; Pelletier, 1977).

Children's difficulties with SI have been variously attributed to general processing limitations (Chierchia et al., 2001; Reinhart, 2004; Pouscoulous et al., 2007), difficulty in understanding the communicative goals of the experimenters (Musolino & Papafragou, 2003; Papafragou & Tantalou, 2004), and a tendency to be more tolerant of pragmatic infelicity (Katsos & Bishop, 2011). More recently, evidence has begun to accumulate in support of an account focusing on children's access to relevant linguistic alternatives that are necessary to derive an SI (Barner et al., 2011; Barner & Bachrach, 2010; Chierchia et al., 2001; Foppolo et al., 2012; Hochstein et al., 2014; Tieu et al., 2016; Skordos & Papafragou, 2016). These accounts argue that difficulties in accessing necessary linguistic alternatives might explain why children fail to derive implicatures, whether such difficulties are due to a lack of associations between scale mates, a failure to detect which alternatives are contextually relevant, or a lack of working memory capacity to compute the contribution of alternatives while simultaneously considering a sentence's basic meaning.

While most studies in support of the "access to alternatives" view have focused on conditions under which children derive weak vs. strong interpretations of utterances, both of which are available to adults, two recent studies have argued for this account by pointing to an entirely different form of evidence, namely the interpretation of disjunction as conjunction. Specifically, these studies report that, given an utterance like (2a) above, some children conclude that the girl must have both types of fruit. In one study documenting this phenomenon, Singh et al. (2016) tested 4- and 5-year-old English-speaking children (N=31) using a modified Truth Value Judgment task (Crain & Thornton, 1998). Surprisingly, they found that children accepted disjunctive statements (e.g., *The boy has an apple or a banana*) only about 35% of the time when one of the disjuncts was true (henceforth 1-Disjunct-True trials), but 76% of the time when both

disjuncts were true (henceforth 2-Disjunct-True-trials). In addition, children displayed a similar behavior with disjunction embedded under a universal quantifier (e.g., *Every boy has an apple or a banana*), accepting 1-Disjunct-True trials about 45% of the time and 2-Disjunct-True trials, 75% of the time. This pattern of responses is particularly surprising when seen from the standard perspective that children often do not derive implicatures at all; if that were the case, they should generally be inclusive and be equally likely to accept disjunction in 1-Disjunct true and 2-Disjunct-True trials. Critically, Singh et al. classified children according to their individual response patterns: Four children performed in an adult-like manner, accepting 1-Disjunct-True utterances while rejecting 2-Disjunct-True utterances (henceforth "Exclusive"). Another four children showed the pattern of failure often reported in previous studies, accepting both 1-Disjunct-True and 2-Disjunct-True trials (henceforth "Inclusive"). Finally, 21 out of 31 children (or 67%) responded as if *or* was *and*, rejecting 1-Disjunct-True utterances while accepting 2-Disjunct-True trials (henceforth "Conjunctive").[1]

Adopting the "access to alternatives" account, Singh et al. argue that children in their study arrive at conjunctive interpretations because they lack access to stronger alternative statements

[1] Against this hypothesis that children simply think *or* always means *and*, Singh and colleagues (2016) point out that children do not derive conjunctive meanings in downward entailing environments, such as when disjunction is embedded under the scope of negation (Gualmini & Crain, 2002; Goro & Akiba, 2004; Jing, Crain & Hsu, 2005), when it occurs in the antecedent of conditionals (Su, 2014), and when it is in the nuclear scope of *only*, *before*, *not ever*y, etc. (Goro, Minai & Crain, 2005; Jing, Crain & Hsu, 2005; Gualmini & Crain, 2002; Notley, Zhou, Jensen & Crain, 2012; cf. Tieu et al., 2017; Singh et al., 2016, for discussion).

derived by replacing *or* with *and.* Consequently, they are restricted to deriving implicatures from alternatives that are contained within the original assertion – i.e., the individual disjuncts. Also, and critical to their account, Singh et al. adopt the grammatical view of implicature (Fox, 2007; Chierchia, et al., 2009), and propose that children's disjunctive alternatives are exhaustified before being negated, resulting in multiple exhaustification (for additional discussion of this account, see Crnič, Chemla, & Fox, 2015; Fox & Katzir, 2011; Franke, 2011; Katzir, 2007; Spector, 2016). Specifically, when children hear a disjunctive statement like (3a), they compute alternatives by exhaustifying each separate disjunct in (3b) and (3c), resulting in the statements in (3d) and (3e):

(3) a. The girl has an apple or an orange.

b. The girl has an apple.

c. The girl has an orange.

d. The girl has an apple but not an orange (i.e., only an apple).

e. The girl has an orange but not an apple (i.e., only an orange).

The original disjunctive statement is then strengthened by negating these exhaustified alternatives, resulting in the statements in (4a) and (4b), which, when combined with the original statement result in a conjunctive meaning, in (4c).

(4) a. It's not the case that the girl has only an apple.

b. It's not the case that the girl has only an orange.

c. The girl has an apple or an orange (6a), but doesn't have only an apple, and doesn't have only an orange (7a & 7b) (i.e., the girl has both an apple and an orange).

According to this theory, adults do not derive the implication in (4c) because they have access to the alternative with *and* – i.e., "The girl has an apple and an orange." The inclusion of

this alternative, since it is stronger than the plain disjunctive statement, leads to its negation, and the corresponding inference that the girl does not have both.[2] This inference directly contradicts the inference in (4c). Since contradictory inferences are not permitted by the exhaustification operator, the derivation of (4c) is blocked (see Fox's 2007 discussion of *innocent exclusion* for details).[3] Thus, on Singh et al.'s analysis, their study simultaneously supports the idea that access to alternatives limits children's implicatures, as well as the idea that children's implicatures can involve multiple iterations of exhaustification. As they note, such an analysis integrates work on acquisition (e.g., Paris 1973; Braine & Rumain 1981; Tieu et al. 2017) with a broader account of multiple exhaustification, which previous studies have used to explain other forms of implicature, including so-called "free choice" inferences (see Fox, 2007; Chierchia et al. 2009; see also Geurts, 2010; Franke 2011; Meyer, forthcoming, for accounts that do not employ multiple exhaustification).

The power of this unifying account depends on the reliability of the developmental data, and Singh, et al.'s findings are not conclusive in this regard. Whereas Singh et al.'s study tested a total of 56 children, only 31 were included in the final sample due to exclusions. In particular, 25 children were excluded because they failed to comprehend *and* as conjunction in control trials.

---

[2] Note that since "The girl has an apple and an orange" is one of the strongest possible utterances, the exhaustification operator has no effect on this alternative. Hence, Exh(⟦ The girl has an apple and an orange ⟧) = ⟦ the girl has an apple and an orange ⟧ .

[3]It is important to note that the inference that the girl doesn't have both is still computed due to the first level of exhaustification. "Exh(Exh(p or q))" minimally asserts "Exh(p or q)," which, if "and" is an alternative, is equivalent to "p or q but not both".

The logic of this exclusion was that this subset of children failed to comprehend the task and therefore could not contribute meaningful data to trials testing children's understanding of *or*. A problem with this logic, however, is that it assumes that one type of failure – failure to comprehend *and* in an adult-like way – is mere noise, while another type of failure – failure to comprehend *or* in an adult-like way – is meaningful. Whereas the experimenters systematically eliminated children with a disjunctive bias for both *and* and *or* (while retaining children with a conjunctive bias), they might have instead excluded all of the children who treated *or* as conjunctive. In that case, they might have found that *and* has a disjunctive meaning. Further, even if we accept the logic of Singh et al. and assume that the difficulty of their task required excluding 25 children, we might worry that task difficulty could also have affected the remaining non-excluded children. Although the 31 children who were included in their study met the relatively modest threshold of comprehending conjunction, the task may have nevertheless been hard enough that it was confusing even to these children, and may not have provided a valid test of their knowledge of logical connectives.

These potential problems with Singh et al. notwithstanding, a similar study by Tieu et al. (2017) found comparable results, while also reporting fewer data exclusions (8 children out of an original 54). In their study, Tieu and colleagues tested 4- and 5-year-old speakers of French and Japanese (N=46) using a Truth Value Judgment task. Rather than asking children to assess descriptions of static scenes as in Singh et al., Tieu et al. instead presented the task as a guessing game, on the rationale that using disjunctive statements to describe visible states of affairs is pragmatically infelicitous and potentially confusing to children. Similar to Singh et al., they found that children accepted 1-Disjunct-True trials only 51% of the time while accepting 2-Disjunct-True trials 81% of the time (with no significant difference between French and

Japanese children). Finally, like Singh et al. (2016), Tieu and colleagues classified children into groups and found that 14 children were Inclusive, 3 were Exclusive, and 19 were Conjunctive (~ 41%). Ten additional children did not fit a predefined pattern, and were classified as "Other".

While the multiple exhaustification hypothesis is elegant in its ability to give a single model-theoretical account for both adults' free-choice readings of disjunction and children's quite different conjunctive readings, there are several reasons to suspect there may be alternative explanations for the developmental findings. First, the evidence for children's conjunctive interpretations of *or* outside of Tieu et al.'s and Singh et al.'s paradigms is limited. Most previous studies were not designed to directly assess conjunctive readings of *or*, with the result that there are relatively few demonstrations of this effect. For instance, some earlier studies did not include 1-Disjunct-True trials (Chierchia et al., 2001; Gualmini et al., 2001), and explicitly included the scalar alternative *and* in the paradigm. Others did not analyze their data in a way that would assess conjunctive interpretations (e.g., Hochstein, Bale, Fox, & Barner, 2014). Further, in one of the previous cases in which evidence of conjunctive responses is reported, the rate of such responses was highly variable and task dependent, making it unclear how robust the finding is both within and across different experimental paradigms (Braine & Rumain, 1981). In another, Paris (1973) reported conjunctive behaviors in children when presented with sentences like, e.g. *The bird is in the nest or the shoe is on the foot*. However, he also found conjunctive behaviors in 8th graders (13-year-olds) and college-age students, who presumably have access to *and* as an alternative to *or*. From these results, Paris did not conclude that children interpret *or* conjunctively due to a semantic operation, but instead that the sentences used may not have been pragmatically felicitous in the contexts used in his study. Finally, one direct replication of Tieu, et al.'s paradigm failed to find a comparable number of conjunctive responders. Sauerland and

Yatsushiro (2018) used the same stimuli as Tieu et al. 2017 with German-speaking children, and found that only 2 out of 31 children in the 4- to 6-year-old age group consistently rejected the 1-Disjunct-True trials while accepting the 2-Disjunct-True trials. This may point to differences between languages, fidelity of methods across different labs, sampling error (since small sample sizes generate a wide range of effect size estimates), or a "decline" effect, wherein replications of early reports exhibit regression to the mean, resulting in smaller effects in later studies (Schooler, 2011).

Related to this first issue regarding the robustness of conjunctive responding is a concern raised by Tieu et al. In their study, they worried that aspects of the Singh et al. study may have proved confusing or infelicitous to children – e.g., that uttering a disjunctive expression like, *The boy is holding an apple or an orange* might be pragmatically odd in a context in which the speaker can see that the boy is holding, e.g., an orange but not an apple. Therefore, the utterance might be rejected because it implies, wrongly, that the speaker is ignorant (for evidence that 5-year-olds infer ignorance of speakers in exactly these situations, see Hochstein, Fox, Bale, & Barner, 2014). Such factors might explain, in part, why the task used by Singh et al. resulted in high rates of exclusion (i.e., because once confused, children performed randomly), and possibly why children provided non-adult interpretations of disjunction.[4] Arguably just as confusing to children, however, are other properties of this design, also featured in the Tieu et al. study. In

[4]Specifically, it might explain why children have difficulty interpreting disjunction when it does not appear within the scope of a universal quantifier. However, since ignorance implicatures are not as readily available within the scope of universals, this can't likely explain the conjunctive behavior of children with respect to these types of sentences.

Tieu et al., for example, children were presented a scene in which a chicken stood atop a hill with two objects (e.g., a bus and an airplane), and a puppet guessed what would happen next in the scene. On critical disjunctive test trials, the puppet guessed, e.g., "The chicken pushed the bus or the airplane", after which the chicken either pushed one of the two objects or both of them down the hill. This occurred in all critical trials, such that for all 1-Disjunct-True trials the character interacted with one of the two available objects and in the 2-Disjunct-True trials it interacted with both. Importantly, it was never the case that the chicken did not act at all (i.e., it always acted on at least one item present in the scene). The problem is that, on all of these trials, the puppet's predictions never carried any actual information, because their guesses could only prove correct (i.e., acting on one or two objects), since not acting at all was never shown to be a possibility. Just as it is odd to predict the outcome of a tennis match, by saying, "*Either A or B will win*," the disjunctive predictions in the Tieu et al. study may have been infelicitous and therefore confusing to children.[5] On the one hand, such a prediction is good, because it is so

[5] Note that there was one control trial where the character acted upon neither of the two objects mentioned in a disjunctive statement (0-Disjunct-True trials). However, in this trial, the character acted on a third alternative that was present in the scene – an alternative that was not present in any of the other trials. Although a basis for rejecting utterances existed within these particular control trials, this was not true of the critical (1-Disjunct-True and 2-Disjunct-True) trials, in which the third object was never present. On these trials, a disjunctive statement could only be false if the character did not perform the mentioned action on either of the two objects in the scene. This is a situation participants never saw, not even on the 0-Disjunct-True control trials. Given this pattern in the experiment, it is possible that both the control and critical trials created

likely to be correct. On the other hand, it's not really a prediction at all, since other alternative outcomes are so unlikely. As we note in the Discussion, this question may also have impacted the results of Singh et al.

In the Tieu et al. study, this type of pragmatic uncertainty may have been compounded by another aspect of their experimental design. After children were presented with each scene — and before they were asked to make a judgment — they were provided with a verbal description of what they had just witnessed. This description explicitly mentioned both items in the scene, either pointing out that the character acted upon one object, but not the other, or that the character acted on both objects. For instance, on trials in which a chicken pushed objects down a hill, children heard the puppet make their prediction, saw the events transpire, and then were told, e.g., "Look, the chicken pushed that!" (e.g., the bus); "She didn't want to break that one so she didn't touch it. So, was Raffie right?" Although this description was intended to remind children of what they had just experienced, it's possible that instead it complicated matters. First, the description may have made the task more difficult by increasing the time between the puppet's disjunctive statement and the child's response, making it more difficult to remember what the puppet had actually said. Second, the description may have made salient the character's failure to act on one of the two objects, leading children to infer that the character really wanted to act on both items. For example, the explanation, "she didn't want to break that one, so she didn't touch it" implies that the character might have wanted to act on the second item but refrained from doing so, and therefore failed to fully fulfill their goal. Such a focus could lead

---

an expectation that the mentioned character (e.g., the chicken) would perform the mentioned action (e.g., pushing) on some object present in the scene.

children to evaluate the sentence “was Raffie right?” as pertaining to whether the chicken carried out their goal, rather than whether the puppet made the correct prediction. Aside from making the puppet’s original statement more difficult to recall, this intervening material may have created confusion regarding the dimension upon which children were expected to make their judgments.

The present study

In the present study, we tested children in three between-subjects conditions. In the first condition, we sought to replicate Tieu et al. (2017) in English-speaking children, in order to confirm that their results were robust, and also not specific to French and Japanese. To do so, we used materials modeled directly after Tieu et al. In the second condition, we used the same materials, but removed the potentially confusing explanation sentence from the Tieu et al. protocol. Finally, in the third condition, we both removed this potentially confusing sentence and also made a key adjustment to the experiment that attempted to render the disjunctive statements felicitous. This trial type is similar to Tieu et al.’s 0-Disjunct-True trials. Specifically, we added a third alternative object that could be acted upon by the character, such that it became possible to imagine a scenario in which a disjunctive guess might be false. For example, in our version of the task, a character upon a hill might have three objects available to push (a bus, a car, and a plane), such that the prediction “The chicken pushed the bus or the car” could conceivably be false (e.g., if the chicken pushed the plane). Under these conditions, we asked whether children continue to exhibit conjunctive judgments, or whether such behaviors are instead an artifact of decision making under pragmatic uncertainty.

Method

Participants

We tested a group of 126 preschoolers between the ages of 4 and 6 years old, who were reported to be typically developing and whose primary language was English. Each child was assigned to one of three conditions. The *Replication* condition included 43 children (4;0 – 5;9, M=5;0), the *Modified Script* condition included 41 children (4;0 – 5;10, M=4;10), the *Three Alternatives* condition included 42 children (4;0 – 5;11, M=5;0). The children were recruited from preschools in San Diego, CA, and Comox, BC, as well as the Fleet Science Center and the Birch Aquarium in San Diego, CA. Three additional children were tested but excluded from the analysis due to experimenter error.

Materials and Procedure

We used an Acceptability Judgment Task closely modeled after the task in Tieu et al. (2017).[6] In all three conditions, children sat in front of a laptop computer and were shown a series of PowerPoint slides, which depicted clipart stimuli. They were introduced to "Raffie" a puppet dragon character who "loves listening to stories" and were told that along with the experimenter and Raffie, they would watch stories about some friendly animals. The experimenter then asked Raffie to guess what would happen next in the story. The children's task was to say whether they thought Raffie did a good job at guessing and was right, or did a bad job at guessing and was wrong. Following the introduction, the experimenter made a mock video-call on the laptop and brought up a small video-window of Raffie to start the experiment. In each trial, children were shown an initial scene depicting an animal character and objects the character would interact with. A short script, read by the experimenter, identified the animal character

[6] We thank the authors for sharing their materials of the original study, which formed the basis for our own stimuli.

(e.g., a ladybug), provided some context (e.g., she loves to paint) and brought attention to what the animal character could possibly do next (e.g., decide what she would like to paint). Then Raffie appeared in her video-window and the experimenter asked the puppet to guess what they thought actually happened in the story. The puppet answered by offering a statement (e.g., "The ladybug painted the cup"). Finally, an outcome scene showed what the animal character actually did and children were asked to judge whether the puppet guessed correctly or not. An example of a training trial can be seen in Fig 1 below.

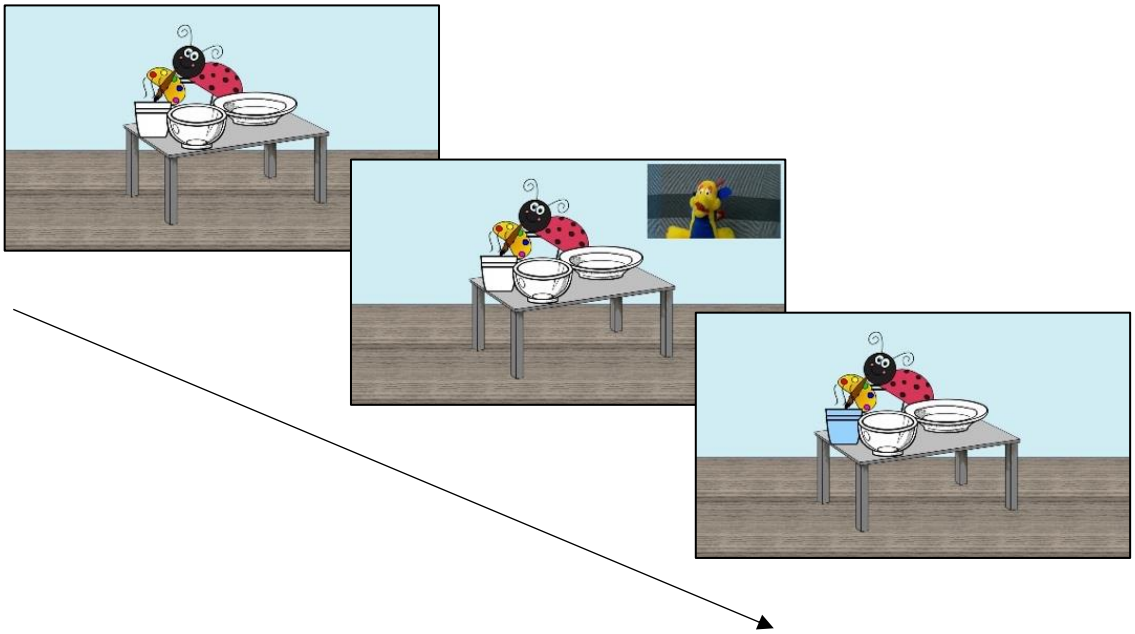

Figure 1. Example of a training trial.

In all conditions, children first completed two training trials. These consisted of scenes like the one in Fig. 1 above, for which Raffie's guess did not include the words *or* or *and*. In one of the training trials, Raffie guessed correctly and therefore she was right, and in the other one

Raffie guessed incorrectly and therefore was wrong. For the training trials children were given feedback.

Following the training trials, the experimenter moved to the test phase of the experiment, which consisted of four 1-Disjunct-True trials, four 2-Disjunct-True trials, two 0-Disjunct-True control trials and three Filler trials, for a total of 13 within-subjects trials, in two orders, each the reverse of the other. For the Filler trials, Raffie made predictions without disjunction, similar to the training trials (e.g., "The pig put on his hat"). Raffie's predictions were true for two of the Filler trials and false for the other one. For all remaining trials, Raffie made predictions using disjunctive statements of the type, "The [animal character] [verbed] an X or a Y". In the 1-Disjunct-True trials the animal character interacted with only one of the objects, in the 2-Disjunct-True trials the animal character interacted with both of the objects, and finally in the 0-Disjunct-True trials the animal character did not interact with any of the mentioned objects, but instead interacted with a third item in the scene, as in the 0-Disjunct-True trials in Tieu et al. (2017). Examples of the final (outcome) scenes of the trials and accompanying statements can be seen in Fig 2 below and full scripts of example trials are available in supplementary materials to this article, available on the journal's website.

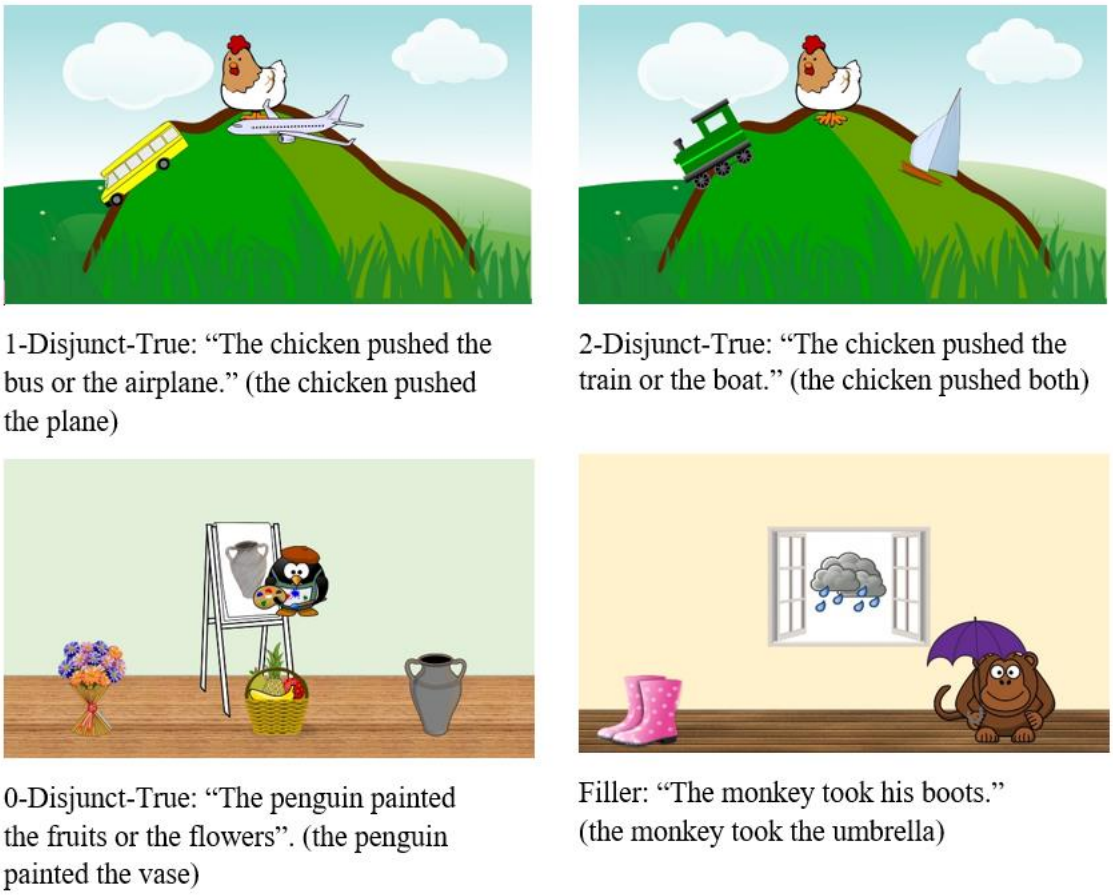

Figure 2. Examples of outcome scenes and accompanying statements of the main phase trials. The actual outcome is described in parentheses.

There were the following differences between the three conditions. First, in only the *Replication* condition, after children watched the animal character act on one or two objects, the experimenter immediately re-described what had happened, which included reference to the objects. For example, in the 1-Disjunct-True trial (Fig. 2), where the puppet (Raffie) said, "The chicken pushed the bus or the airplane" but only pushed the airplane, the experimenter pointed to the bus and said, "Look, the chicken pushed that!" (i.e., the bus). "She didn't want to break that one (i.e., the airplane) so she didn't touch it. So was Raffie right?" It is possible that this

reminder description, which inserted text between the prediction and the child's assessment, may have made it more difficult for the child to recall the original utterance, particularly after many repeated trials with highly similar utterances and outcomes. Therefore, this reminder was removed in the *Modified Script* condition, which was identical to the *Replication* condition otherwise. Finally, the *Three Alternatives* condition was very similar to the *Modified Script* condition apart from the following difference: All trials which featured two objects in the *Modified Script* (and in the *Replication*) condition now included a third object, as in the 0-Disjunct True condition of Tieu et al. (see Figure 3).

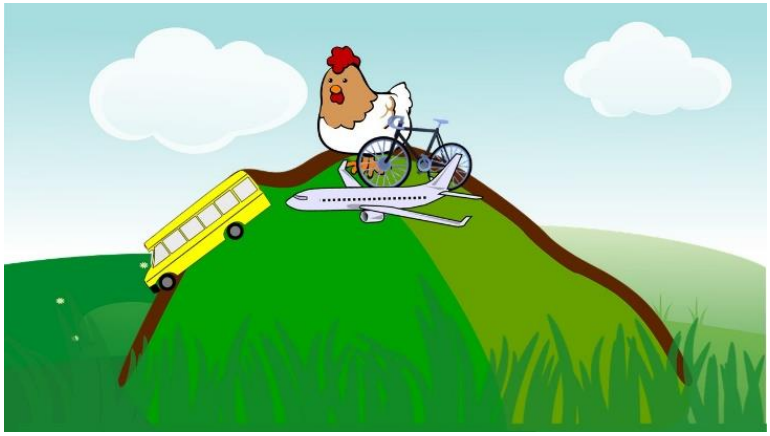

Figure 3. 1-Disjunct-True (*Three Alternatives* condition). Notice the addition of a third item (bicycle) in the scene with the chicken.

Results

We first describe a statistical model that compares children's performance across the three experimental conditions, to assess the degree to which they accepted 1-Disjunct-True and 2-Disjunct-True trials across the two conditions. Next, we categorized children according to their

individual response patterns of accepting and rejecting trial types, following the scheme used by Tieu and colleagues (2017) to examine how many children fall into the "Conjunctive" pattern of responders and how a more pragmatically felicitous disjunctive prediction in the *Three Alternatives* condition may have affected children's generalizations.

Overall, children correctly accepted the True Fillers (*Replication:* 91% correct; *Modified Script:* 90%; *Three Alternatives:* 89%; $p = 0.939$[7]) and rejected both the 0-Disjunct-True trials (*Replication:* 79% correct; *Modified Script:* 79%; *Three Alternatives:* 80%; $p = 0.983$) and the False Fillers (*Replication:* 84% correct; *Modified Script:*84%; *Three Alternatives:* 80%; $p = 0.998$) across conditions. In sum, performance on control items did not differ across conditions. Of particular interest were children's responses to the 1-Disjunct-True and 2-Disjunct-True trials. Following Tieu et al., (2017) we constructed separate mixed-effects logistic regressions for 1-Disjunct-True and 2-Disjunct-True trials with Condition (*Replication, Modified Script, Three Alternatives*) as a fixed factor and intercepts for Subject and Item as random effects. For 2-Disjunct-True trials there was no effect of Condition ($M_{\text{REPLICATION}}$=.90; $M_{\text{MODIFIED SCRIPT}}$=.80; $M_{\text{THREE ALTERNATIVES}}$=.88; $\chi^2(2) = 3.721$, $p = .156$) with children being equally likely to accept them across the three conditions. For 1-Disjunct-True trials there is an effect of Condition ($M_{\text{REPLICATION}}$=.58; $M_{\text{MODIFIED SCRIPT}}$=.66; $M_{\text{THREE ALTERNATIVES}}$=.78; $\chi^2(2) = 7.717$, $p = .021$), with children in the *Three Alternatives* condition being significantly more likely to accept 1-Disjunct-True trials than children in the *Replication* condition ($z = 2.720$, $p = 0.007$). Children in

[7] These comparisons represent separate GLMM analyses for each trial type (True Fillers, False Fillers, 0-Disjunct-True trials) with Condition (*Replication, Modified Script, Three Alternatives*) as a fixed factor and intercepts for Subject and Item as random effects.

the *Modified Script* condition accepted 1-Disjunct-True trials at a rate that was intermediate to the other two conditions, and which did not differ statistically from either the *Replication* condition ($z = 1.015$, $p = 0.310$), or the *Three Alternatives* condition ($z = -1.826$, $p = 0.068$). This suggests that simply removing the reminder description in the *Modified Script* condition did not alone have an effect, but that removing both the description and adding a third alternative in the *Three Alternatives* condition made children significantly more likely to accept the 1-Disjunct-True trials.

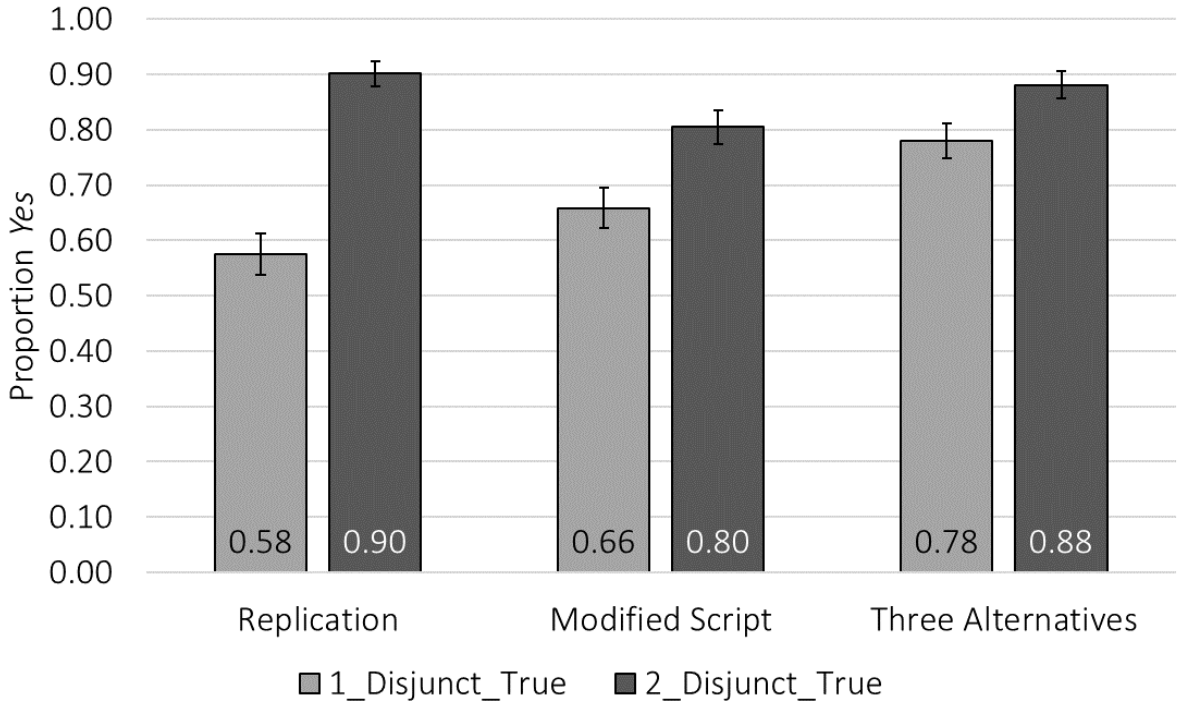

Figure 4. Children's acceptance of disjunctive statements averaged across trials and subjects.

To determine how many children in our study exhibited a conjunctive pattern of responding, we next classified children following the scheme of Tieu et al. (2017). Specifically, if children accepted at least 3/4 1-Disjunct-True trials and rejected at least 3/4 2-Disjunct-True trials they were classified as "Exclusive". Children who accepted at least 3/4 1-Disjunct-True and 2-

Disjunct-True trials were classified as “Inclusive”. Children who rejected at least 3/4 1-Disjunct-True and accepted at least 3/4 2-Disjunct-True trials were classified as “Conjunctive””. All other children were classified as “Other”. Table 1 shows the results of this classification scheme in each condition.

In the *Replication* condition, there were 21 Inclusive children, 1 Exclusive child, 14 Conjunctive children, and 7 children classified as “Other”. These numbers closely resemble Tieu et al. (2017), who found 14 Inclusive, 3 Exclusive, 19 Conjunctive, and 10 Other.[8] Although there were only 7 Conjunctive children in the *Modified Script* condition (21 Inclusive, 4 Exclusive, 9 Other), a Fisher’s Exact Test on a 2x2 contingency table comparing the number of Inclusive vs. Conjunctive children across the two conditions found no statistically significant difference (p = .284). In contrast, results were substantially different in the *Three Alternatives* condition. Here, the number of Inclusive children increased to 31 while the number of Conjunctive children was only 4. There was again only 1 Exclusive child and 6 children who were classified as “Other”. A Fisher’s Exact Test comparing the number of Inclusive and Conjunctive children between the *Replication* and *Three Alternatives* condition revealed a significant difference between the two, reflecting the fact that the *Replication* condition featured

[8] We did not exclude children for failing control trials as in Tieu et al. (since random guessing predicts equal numbers of correct and incorrect guesses, and it is impossible to symmetrically identify and remove correct guesses; see our discussion of data removal in the Introduction). Removing these children from our analyses has the following effects: a) It reduces the number of Inclusive children across the three conditions and b) returns a statistically significant difference between the *Modified Script* and *Three Alternatives* conditions on 1-Disjunct-True trials.

significantly more Conjunctive and fewer Inclusive children compared to the *Three Alternatives* condition (one-tailed $p = 0.006$). Critical to assessing the significance of the 4 remaining Conjunctive children identified in our study, this was fewer than the 6 children classified as "Other" – a category which Tieu, et al. took to indicate noise. As we discuss in the General Discussion, some number of Conjunctive children should always be expected if some proportion of children guess randomly. With fewer Conjunctive than Other responders, random guessing may be the simplest explanation.

Table 1. Individual response patterns by condition.

	Inclusive	Exclusive	Conjunctive	Other
Tieu et al., 2017	14	3	19	10
Replication	21	1	14	7
Modified Script	21	4	7	9
Three Alternatives	31	1	4	6

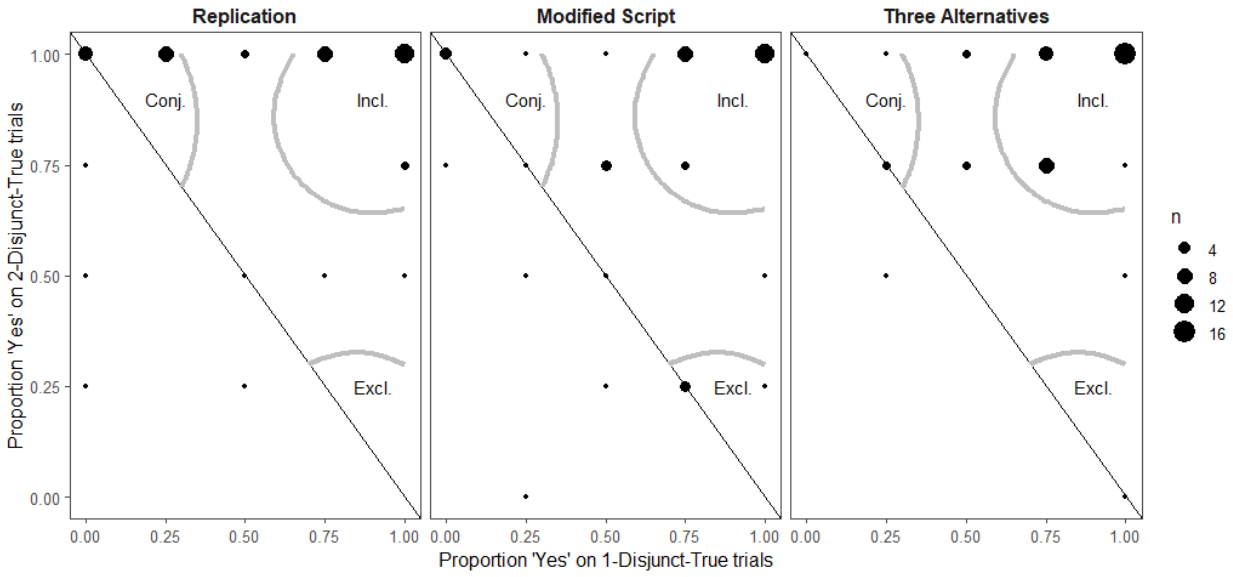

Figure 5. Children's rate of accepting disjunctive predictions on 1-Disjunct-True and 2-Disjunct-True trials, modeled after Figure 3 in Tieu, et al. (2017). Vertices represent possible combinations of acceptance rates of each trial for one individual child. The size of each point corresponds to the number of children who responded at these rates. The top left corner corresponds to 'conjunctive' responders: children who rejected 1-Disjunct-True and accepted only 2-Disjunct-True trials. The bottom right corner corresponds to 'exclusive' responders: children who accepted only 1-Disjunct-True trials, and 2-Disjunct-True trials. The top right corner corresponds to 'inclusive' responders: children who accepted disjunctions in either case.

Discussion

In this study, we investigated the recent claim that a substantial percentage (41% - 67%) of preschool aged children interpret disjunctive statements of the form *x or y* as conjunctive (Singh et al., 2016; Tieu et al., 2017). In particular, we explored the hypothesis that children who appear to interpret disjunction conjunctively might be confused about the intended question of the experimenter. To test this, we first replicated Tieu and colleagues (2017) in English-speaking children in our *Replication* condition. When tested in this paradigm, we found that a substantial subset of children (14/43; 33%) performed as though "conjunctive" and accepted disjunctive statements as descriptions of scenes in which both disjuncts were true, but not of scenes in which only one of the disjuncts was true. In the *Modified Script* condition, in which we removed a potentially confusing sentence from the script of the experiment 7 of 41 children (17%) performed as though conjunctive. Finally, in our *Three Alternatives* condition, where the puppet's predictions could plausibly be false, we found almost no conjunctive responders (4/42; 9.5%). While we cannot be sure whether the addition of a third alternative would alone be sufficient to reduce conjunctive responding, or whether the third alternative only had an effect in

combination with the removal of the reminder statement, the key conclusion of this pattern of findings is that these contextual manipulations reduced the number of conjunctive children to almost zero, and below the number of children who did not interpret *or* in any systematic way and were classified as "Other".

These data lead us to two main conclusions. First, in keeping with a recent failure to replicate Tieu et al. (Sauerland et al., 2017), our data suggest that the likelihood of conjunctive responding may be smaller than estimated by initial reports of this phenomenon, and may therefore not be detected in all studies. Second, our data also lead us to question the *nature* of this effect, and whether it reflects semantic competence or instead is the result of pragmatic uncertainty regarding the experimenter's intended question (as proposed in early reports of conjunctive responding, like Paris, 1973, who found it not only for disjunction, but also for conditional statements).[9] As noted in the Introduction, some children may have been confused by descriptions that intervened between the primary stimuli and the child's own judgment, e.g., either because the descriptions were themselves confusing, or because their addition made it more difficult for children to recall which prediction the puppet made (since multiple predictions were made in a single short session). Compatible with this concern, only 7 children remained

[9] In his study, Paris (1973) found that children not only accepted disjunctive statements when both disjuncts were true, but also that they accepted conditionals only when both the antecedent and the consequent were true. Presented with a sentence like, "If the bird is in the nest, then the shoe is on the foot" children judged the sentence as true only when both the bird was in the nest and the shoe was on the foot (for similar results, see Rumain et al., 1983, Barrouillet, Gauffroy, & Lecas, 2008; Gauffroy & Barrouillet, 2009; Markovits et al., 2016).

conjunctive when this extraneous material was removed. Relatedly, we worried that when children are tested with 2 alternatives, some may be confused by the fact that the puppet's guess carried no information regarding the future acts of the chicken, and may therefore infer that the question cannot relate to the truth of the disjunctive statement (since it is necessarily true under any outcome where the chicken pushes something), but to some other question under discussion (QUD) that makes sense with respect to the situation at hand. Such an analysis is compatible with Crain et al.'s notion of "plausible dissent", which argues that when children are unable to easily imagine a context in which an utterance might prove false, they evaluate it using idiosyncratic confabulated criteria (e.g., a contextually relevant alternative QUD). Adding a third alternative provides a condition in which the utterance can be false, satisfying plausible dissent, and leading to virtually no remaining conjunctive responses.[10]

Critically, the hypothesis that conjunctive responding is the product of confusion is not itself meant to be a formal account of what generates conjunctive responses. Instead, our contention is that not all behavioral responses signal systematic knowledge states that merit semantic modeling. First, children often provide behavioral responses which we know are at odds with their cognitive states, either due to problems of inhibitory control (Zelazo et al., 1995; Munakata et al., 2011), or pragmatic infelicities of experimental design. For example, Piaget famously claimed that 5-year-old children lack a complete concept of number because, after counting two lines of objects and judging them to be equal, these children can be led to judge that one of these

[10] This analysis can also explain Singh et al. (2016), in which disjunctive statements were also always trivially true (i.e., there was no basis provided in the context that might render them false).

lines has *more* objects than the other by the experimenter spreading it out. Subsequent work showed that these responses had nothing to do with number word knowledge at all, but instead children's uncertainty about what question was intended by the experimenter. In one such study, McGarrigle and Donaldson (1978) showed that if one line of objects is stretched *accidentally* rather than intentionally (i.e., by an angry teddy), children no longer interpret the transformation as relevant to the speaker's request and continue to judge the two sets as equal when asked a second time. Likewise, if children are simply asked the question only once – after the transformation – they judge the two sets to be numerically equal (Rose & Blank, 1974). These studies show that five-year-old children know that 1-to-1 correspondence establishes that two sets have the same cardinality, and that length is irrelevant to equality, but that they are willing to override their semantic knowledge and consider length if this renders the experimenter's question felicitous in the context. Piaget's original conclusion that children lack knowledge of number is unwarranted, even if there is no alternative formal account of children's failure on the original task.

Second, although it is potentially meaningful when children provide consistent responses, children notoriously perseverate in ways that conflict with knowledge they reveal in different paradigms, or in ways that are entirely meaningless. For example, although 3- and 4-year-old children can easily sort objects by either shape or color, when they're asked to switch dimensions halfway through a study, they often fail and continue sorting the same way, despite the fact that this results in responses that conflict with their independently demonstrated category knowledge (Zelazo et al., 1995). More strikingly, children also perseverate when their initial response does not reflect any particular knowledge state (Osler & Kofsky, 1965). For example, when 4- to 8-year-old children are shown shapes that vary in color and size and are asked to guess whether

they belong to a category (e.g., where "red squares" might be rewarded), half of the children who fail to learn the correct category repeatedly make the same response, even though they are given feedback that this response is incorrect. Findings such as these indicate that children often respond consistently in an experiment even when these responses are contrary to their knowledge state, or follow an initial random guess.[11]

Our data suggest that conjunctive responses are quite infrequent, and even without manipulating meaningful aspects of study design – e.g., the number of alternatives – conjunctive responders represent only about 17% of all children. Although we believe that this pattern of findings is unlikely to be the systematic product of multiple exhaustification or any other single factor, one might consider making additional assumptions in an effort to explain this small segment of the data via a formal mechanism. For example, to explain the reduction in conjunctive responders in our Three Alternatives condition, one might hypothesize that adding extra objects expands the range of alternatives considered, and as a result changes which parses the listener prefers. In their paper, Singh et al. (2016) propose that parses which provide

[11] In past work, some children provide justifications for conjunctive responses that are potentially compatible with a conjunctive semantics. For example, when asked why they rejected disjunctive sentences when only one disjunct was true some responded by noting that one of the disjuncts was false (e.g., *he doesn't have both*). However, because these statements are just true descriptions of scenes, and not necessarily explanations of why the scenes led children to their judgments, they only tell us that children know what happened, not how they interpreted the utterance.

complete answers to a QUD[12] are preferred over those that do not, and that if there is no parse that provides a complete answer, simpler parses without exhaustification should be preferred instead. In particular, they assume that children have three different parses available to them when interpreting disjunctive statements: One without any exhaustification operators, which results in an inclusive interpretation, one with only one exhaustification operator, which also results in an inclusive interpretation, and one with two exhaustification operators, which results in a conjunctive interpretation.[13]

This hypothesized parsing strategy predicts a difference between 2-object and 3-object trials. For example, if we assume that the inferred QUD in the experiment is something like, "What does the chicken push down the hill?", then what counts as a "complete answer" changes with respect to the number of objects in the scene. In the 2-object trials (where only a bus and an airplane are present), the conjunctive parse of "The chicken pushes the bus or the airplane" would provide a complete answer (and a more informative prediction). Such a parse would inform the hearer about what happens to each object (i.e., the chicken pushes both). In contrast, the other parses would only partially answer the QUD—the parses would give the hearer some information (i.e., that at least one of the objects was going to be pushed) but ultimately not a complete picture relative to the QUD (i.e., the inclusive parses would not inform the hearer about

[12] Singh et al. define QUD in the formal sense as outlined in Groenendijk and Stokhof (1984). We are using the term in the same sense. See also the discussions in Lewis (1988), Roberts (1996), and references therein.

[13] Note that the addition of any further exhaustification operators results in a parse that is equivalent to the option with two operators.

which of the objects were going to be pushed nor about whether only one or two objects were going to be pushed). When a third object is present, however, the situation changes. Now, the conjunctive parse fails to address the outcome of the third object, and hence no longer provides a complete answer to the QUD. The conjunctive parse provides some information (i.e., what happens to two of the objects—the bus and the airplane) but it does not provide a complete picture (i.e., the hearer has no information about whether or not the chicken pushes the third object—the bike). Therefore, according to Singh et al.'s parsing strategy, the conjunctive parse should no longer be preferred over an inclusive parse. In fact, since the inclusive parses are syntactically simpler, they should be preferred over the conjunctive parse.[14]

However, even this modification may not provide an adequate description of remaining conjunctive responses.[15] A key difficulty is that this proposal only explains our data if we assume

[14] A variant of this idea is to propose that children prefer informative parses over uninformative ones (cf. Huang & Crain, 2019). If children assume that at least one object will be acted upon, then in the 2-object case only the conjunctive parse is informative, among the parses that do not use the 'and' alternative, whereas all three parses provide at least some information in the 3-object context. The result is that in the 3-object case, there shouldn't be any preference for one parse over another, contrary to our experimental results where it is clear that the less-informative inclusive parse is preferred.

[15] In the main text, we outline only one significant challenge but there are others. For example, adults do not always favor simpler parses over more complex ones when no parse is a complete answer to the QUD (i.e., in certain conditions where both inclusive and exclusive interpretations of disjunction are under-informative, adults still prefer an exclusive interpretation over an

that children do not compute an *ad hoc* implicature in the Three Alternatives condition (i.e., if we assume they do not get the implication that unmentioned objects are not acted upon). If, in the Three Alternatives condition, children *do* compute an *ad hoc* implicature, and infer that the unmentioned object was not acted upon (in the example above, that the chicken does not push the bike), then the conjunctive parse resulting from double exhaustification should once again provide a complete answer to the QUD – i.e., it should communicate that the chicken pushes the airplane and the bus but not the (unmentioned) bike. This prediction remains no matter how *ad hoc* implicatures are computed, whether the third object alters the alternatives that are considered by the exhaustification operator (e.g., "The chicken pushed the bike" is an alternative which has the same status as "The chicken pushed the airplane") or whether instead *ad hoc* implicatures are computed by general Gricean mechanisms. Children's ability to compute such *ad hoc* implicatures is well attested (see Barner, Hochstein, Rubenson, & Bale, 2018; Barner et al. 2011; Papafragou & Tantalou, 2004; Stiller et al. 2015).[16]

---

inclusive one despite the fact that the exclusive parse is plausibly more complex: see Huang & Crain under review). Also, such a parsing strategy has difficulty explaining why conjunctive parses would be completely absent in conditions that test for ignorance implicatures (see Hochstein et al. 2014).

[16] Note that the alternative proposal outlined in footnote 12 is unaffected whether or not ad hoc implicatures are posited, since in either case both an inclusive and a conjunctive parse should be considered when 3 objects are present, whereas only the conjunctive parse should be considered in a 2-object context. However, on that proposal there is no explanation for why children shouldn't continue to prefer a conjunctive parse, since it provides more information than an

Although our data bear most directly on the role of multiple exhaustification in child language, they may also have consequences for hypotheses regarding adult language. Most extant data testing children's interpretations of disjunction, including ours, can be explained by single exhaustification and a lack of access to conjunctive alternatives, similar to what has been argued for scalar implicatures involving quantifiers (Barner et al., 2011; Barner & Bachrach, 2010; Chierchia et al., 2001; Foppolo et al., 2012; Hochstein et al., 2014; Horowitz, Schneider, & Frank, 2017; Skordos & Papafragou, 2016; Stiller, Goodman & Frank, 2015). This is important, because multiple exhaustification is often invoked to explain free choice readings (Kratzer & Shimoyama, 2002; Fox, 2007; Alonso-Ovalle, 2006, Chierchia, 2006), in which children and adults derive conjunctive interpretations for utterances like "The girl is allowed to eat pizza or pasta for lunch" (Tieu et al., 2015; Zhou et al., 2013; Pagliarini et al., 2018, for discussion). On such accounts, free choice inferences arise from multiple exhaustification and do not consider conjunctive alternatives, and therefore are computed much like the conjunctive inferences proposed by Tieu et al. (2016) and Singh et al. (2016).[17] If children do not deploy multiple exhaustification to interpret disjunctive statements, or do so very infrequently, then some other mechanism might be required to explain children's systematic conjunctive readings of free-choice statements, which they compute over 90% of the time (Tieu et al., 2015; Zhou et al., 2013). One possibility, for example, are accounts that invoke semantic mechanisms available

inclusive one even in a 3-object context being the only parse that provides a complete answer that addresses the outcome for all three contextually available objects.

[17] Note that all accounts assume that children lack access to conjunctive alternatives when interpreting disjunction, as shown, for example, by their failure to compute exclusivity inferences. Assuming that children *can* access stronger alternatives would not rescue multiple exhaustification, since it would then predict exclusivity inferences in children, as in adults.

to both adults and children, such as those that build free-choice inferences into the semantic interaction between modals and disjunction, separate from the mechanisms that are used to derive scalar implicatures (see Aloni, 2003, 2007; Bar-Lev & Fox, 2017, Klinedinst, 2007; Simon, 2005; Zimmermann, 2000). Though evidence from children alone should not be decisive in deciding between theories, and many other considerations are relevant (Crnič et al., 2015; Fox & Katzir, 2011; Franke, 2011; Katzir, 2007; Spector, 2016), our findings nevertheless argue that child data do not provide straightforward evidence in favor of multiple exhaustification. Future studies should explore this issue, as well as the possibility that multiple exhaustification emerges later in acquisition, or that alternative versions of multiple exhaustification (e.g., that don't predict conjunctive children) might better conform to children's behaviors.

References

Aloni, Maria. 2003. Free Choice in modal contexts. In M. Weisgerber (ed.), Sinn und Bedeutung 7, 1–13. Universität Konstanz.

Aloni, Maria. 2007. Free Choice, modals, and imperatives. Natural Language Semantics 15(1). 65–94.

Alonso-Ovalle, L. (2006). *Disjunction in alternative semantics*. UMass Amherst PhD dissertation.

Bar-Lev, M. E., & Fox, D. (2017). Universal free choice and innocent inclusion. *Semantics and Linguistic Theory*, *27*, 95-115.

Barner, D., & Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology*, *60(1)*, 40-62.

Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition*, *188*, 87-96.

Barner, D., Hochstein, L. K., Rubenson, M. P., & Bale, A. (2018). Four-year-old children compute scalar implicatures in absence of epistemic reasoning. *Semantics in Language Acquisition*, *24*, 325.

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

Brooks, P.J., & Sekerina, I. (2005). Shortcuts to quantifier interpretation in children and adults. *Language Acquisition, 13*, 177-206.

Chierchia, G., (2004). Scalar implicatures, polarity phenomena and the syntax / pragmatics interface. In: Belletti, A. (Ed.), *Structures and Beyond*, (pp. 39-103). Oxford University Press, Oxford.

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the "logicality" of language. *Linguistic Inquiry 37*(4). 535–590. http://dx.doi.org/10.1162/ling.2006.37.4.535.

Chierchia, G., Crain, S., Guasti, M., T., Gualmini, A., & Meroni L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. *Proceedings from the Annual Boston University Conference on Language Development, 25*, 157-168.

Chierchia, G., Fox, D., & Spector, B. (2009). Hurford's Constraint and the Theory of Scalar Implicatures. In: P. Egré & G. Magri (Eds.), *Presuppositions and Implicatures*, (pp. 47–62). Proceedings of the MIT-Paris Workshop. Cambridge, MA: MIT Working Papers in Linguistics.

Chierchia, G., Guasti, M. T., Gualmini, A., Meroni, L., Crain, S., & Foppolo, F. (2004). Semantic and pragmatic competence in children's and adults' comprehension of or. In: Ira Noveck & Dan Sperber, (eds.), *Experimental pragmatics* (pp. 283-300). New York: Palgrave Macmillan.

Crain, S., & Khlentzos, D. (2010). The logic instinct. *Mind & Language*, *25*(1), 30-65.

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

Crnič, L., Chemla, E., & Fox, D. (2015). Scalar implicatures of embedded disjunction. *Natural Language Semantics*, *23*(4), 271-305.

Drozd, K. F. (2001). Children's Weak Interpretations of Universally Quantified Questions. In M. Bowerman and S. C. Levinson, (eds.), *Language Acquisition and Conceptual Development,* Cambridge University Press, Cambridge, England.

Drozd, K. & van Loosbroek, E. (1999). Weak quantification, plausible dissent, and the development of children's pragmatic competence. In A. Greenhill, H. Littlefield & C. Tano (Eds.), *The Proceedings of the 23d Annual Boston University Conference on Language Development* (pp. 184-195). Somerville MA: Cascadilla Press.

Foppolo, F., Guasti, M.T., & Chierchia, G. (2012). Scalar implicatures in child language: Give children a chance. *Language Learning and Development, 8*, 365-394. Psychology Press.

Fox, D. (2007), Free choice and the theory of scalar implicatures. In Uli Sauerland and Penka Stateva (eds.), *Presupposition and implicature in compositional semantics*, Palgrave Macmillan, Houndmills, Basingstoke.

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

Franke, M. (2011). Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics and Pragmatics* 4(1): 1–82. doi:10.3765/sp.4.1.

Freeman, N., Sinha, C. & Stedmon, J. (1982). All the cars – which cars? From word meaning to discourse analysis. In M. Beveridge (Ed.), *Children Thinking through Language* (pp. 53-74). London: Edward Arnold.

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

Geurts, B. (2003). Quantifying kids. *Language Acquisition, 11*, 197–218.

Geurts, Bart (2010). Quantity Implicatures. Cambridge: Cambridge University Press.

Grice, P. (1989). *Studies in the Way of Words*. Harvard University Press. Cambridge.

Gualmini, A., Crain, S., Meroni, L., Chierchia, G., & Guasti, M., T. (2001). At the semantics/pragmatics interface in child language. *Proceedings of Semantics and Linguistic theory XI*. Ithaca, NY: CLC Publications, Department of Linguistics, Cornell University.

Hochstein, L., Bale, A. Fox, D., Barner D. (2014). Ignorance and inference: do problems with Gricean epistemic reasoning explain children's difficulty with scalar implicature? *Journal of Semantics, 33: 107–35*.

Horowitz, A. C., Schneider, R. M., & Frank, M. C. (2017). The Trouble With Quantifiers: Exploring Children's Deficits in Scalar Implicature. *Child development*.

Huang, H., & Crain, S. (under review). When OR is assigned a conjunctive inference in child language.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child: Classification and seriation.* London: Routledge and Kegen Paul.

Katsos, N., & 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: 669–690.

Klinedinst, N. (2007). Plurality and Possibility: University of California, Los Angeles dissertation.

Kratzer, A. & Junko Shimoyama. (2002). Indeterminate pronouns: The view from Japanese. In Yukio Otsu (ed.), Third Tokyo conference on psycholinguistics, 1–25. Tokyo: Hituzi Syobo.

McCawley, J. D. (1993). Everything that Linguists have Always Wanted to Know About Logic* (But Were Ashamed to Ask), 2nd edn., Chicago University Press.

Meyer, M.-C. (Forthcoming). “An apple or a pear: free choice disjunction.” in L. Matthewson, C. Meier, H. Rullmann and T. E. Zimmermann (eds.), The companion to semantics, Wiley’s Linguistics Companion Series.

Noveck, Ira. (2001). When children are more logical than adults: Investigationsof scalar implicature. *Cognition 78:* 165–188.

Osler, S. F., & Kofsky, E. (1965). Stimulus uncertainty as a variable in the development of conceptual ability. *Journal of Experimental Child Psychology, 2*(3), 264-279.

Pagliarini, E., Bill, C., Romoli, J., Tieu, L., & Crain, S. (2018). On children’s variable success with scalar inferences: Insights from disjunction in the scope of a universal quantifier. *Cognition, 178,* 178-192.

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. G. (1973). Comprehension of language connectives and propositional logical relationships. *Journal of experimental child psychology*, *16*(2), 278-291.

Pelletier, F. J.: 1977, *‘Or’, Theoretical Linguistics* 4, 61–74.

Philip, W. (1991). *Spreading in the Acquisition of Universal Quantifiers.* WCCFL 10: 359-373.

Philip, W. (1995). *Event quantification in the acquisition of universal quantification*, Doctoral dissertation. University of Massachusetts, Amherst.

Pouscoulous, N., Noveck, I., Politzer, G., Bastide, A. (2007). Processing costs and implicature development. *Language Acquisition, 14*, 347-375.

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

Roeper, T. & De Villiers, J. (1993). Emergence of bound variable structures. In E. Reuland & W. Abraham (Eds.), *Knowledge and Language: From Orwell's Problem to Plato's Problem* (pp. 105–139). Dordrecht: Kluwer.

Rose, S.A.; Blank, M. (1974). "The potency of context in children's cognition: An illustration through conservation". *Child Development, 45* (2): 499–502.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy, 27*, 367–91.

Sauerland, U., & Yatsushiro, K. (2017). The acquisition of disjunctions: Evidence from German children. *Proceedings of Sinn und Bedeutung 21*.

Schooler, J (2011). "Unpublished results hide the decline effect". *Nature. 470* (7335): 437.

Simons, Mandy. 2005. Dividing things up: The semantics of 'or' and the modal/'or' interaction. Natural Language Semantics 13(3). 217–316.

Singh, R., Wexler, K., Astle-Rahim, A., Kamawar, D., & Fox, D. (2016). Children interpret disjunction as conjunction: Consequences for theories of implicature and child development. *Natural Language Semantics, 24(4): 305-352.*

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

Spector, B. (2003), 'Scalar implicatures: Exhaustivity and Gricean reasoning'. *Proceedings of the ESSLLI 3:* 277–88.

Spector, B. (2006), *Aspects de la pragmatique des ope´rateurs logiques.* Ph.D. thesis, Universite´ Paris 7, Paris

Spector, B. (2016). Comparing exhaustivity operators. *Semantics and Pragmatics, 9*, 11-1.

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

Tieu, L., Romoli, J., Zhou, P., & Crain, S. (2016). Children's knowledge of free choice inferences and scalar implicatures. *Journal of Semantics, 33(2)*, 269–298. DOI:10.1093/jos/ffv001

Tieu, L., Yatsushiro, K., Cremers, A., Romoli, J., Sauerland, U., & 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.*

Zimmermann, T. Ede. 2000. Free Choice disjunction and epistemic possibility. Natural Language Semantics 8(4). 255–290.

APPENDIX A:

The scripts for the same test trial across the three conditions.

Replication Condition

[1] Once there was a chicken who loved to play with her toys, and she especially loved to push them! One day her father gave her two new toys: a bus and an airplane! The chicken was very happy to play with them. Let's see if Raffie can guess what happened next.

[2] EXP: Raffie, tell us, what happened next?

PUP: The chicken pushed the bus or the airplane.

EXP: Let's see if Raffie was right!

[3] Look, the chicken pushed that! She didn't want to break that one so she didn't touch it. So was Raffie right?

Modified Script Condition

[1] Once there was a chicken who loved to play with her toys, and she especially loved to push them! One day her father gave her two new toys: a bus and an airplane! The chicken was very happy to play with them. Let's see if Raffie can guess what happened next.

[2] EXP: Raffie, tell us, what happened next?

PUP: The chicken pushed the bus or the airplane.

EXP: Let's see if Raffie was right!

[3] So was Raffie right?

Three Alternatives Condition

[1] Once there was a chicken who loved to play with her toys, and she especially loved to push them! One day her father gave her three new toys: a bus, a bike, and an airplane! The chicken was very happy to play with them. Let's see if Rafie can guess what happened next.

[2] EXP: Rafie, tell us, what happened next?

PUP: The chicken pushed the bus or the airplane.

EXP: Let's see if Rafie was right!

[3] So Rafie right?