Experimental perspectives on the semantics and pragmatics of plurality

Lyn Tieu   Jacopo Romoli
Macquarie University   *Ulster University*

1   Introduction

Plural morphology in English is associated with a 'more-than-one' meaning. For example, a sentence like (1a), in contrast to its singular counterpart in (1b), is generally interpreted as conveying (2). This data point in itself is unproblematic: on the basis of our intuitions about (1a), we might simply conclude that the literal meaning contributed by plural morphology precisely excludes singularities among the possible denotations of the predicate, as proposed for instance by Chierchia (1998).

(1)   a.   Emily fed giraffes.
       b.   Emily fed a giraffe.

(2)   Emily fed more than one giraffe.

The well-known puzzle associated with these data, however, is the following: if (1a) and (2) share the same meaning, we should expect the negation of the former to be equivalent to the negation of the latter. But this is not the case; the negation of (1a), in (3a), is clearly not equivalent to the negation of (2), in (4). Rather, (3a) conveys what the negation of its singular counterpart in (3b) conveys, namely that Emily didn't feed any giraffe.[1]

(3)   a.   Emily didn't feed giraffes.
       b.   Emily didn't feed a (single) giraffe.

(4)   Emily didn't feed more than one giraffe.

---

[1] There is a marked reading of (3a) that makes it equivalent to (4), as supported by the continuation in (i). This generally requires stress on the syllable containing the *-s* morpheme. We will return to this reading.

(i)   Emily didn't feed giraffes, she fed only one!

The puzzle extends beyond negation, and includes questions such as (5), to which we would intuitively respond affirmatively if Emily fed a single giraffe, in contrast to how we would respond to the question in (6).

(5)   Did Emily feed giraffes?

(6)   Did Emily feed more than one giraffe?

There are two main kinds of approaches to this puzzle in the literature. One approach argues for an ambiguity in the meaning of the plural, and invokes a principle for selecting between the two possible meanings in different contexts (Farkas & de Swart 2010, Grimm 2013). An alternative approach posits that the 'more than one' meaning of plural morphology comes about as a pragmatic inference; specifically, the multiplicity inference is proposed to arise as an *implicature* (Sauerland 2003, 2005, Spector 2007, Zweig 2009, Ivlieva 2013, Magri 2014, Mayr 2015).

Most experimental work on plurality to date has focused on testing some version of the implicature-based approach. These studies have typically focused on four aspects of the multiplicity inference: (i) its sensitivity to logical polarity, (ii) its context dependence, (iii) the possibility of its suspension, and (iv) its development in young children. As we will see, while experimental data from adults regarding (i) and (ii) suggest that multiplicity is indeed a scalar implicature, the existing evidence pertaining to (iii), at least *prima facie*, provides an argument in the opposite direction. We will argue that the *developmental* data may provide the key evidence for adjudicating between the two theoretical approaches. In particular, data obtained from preschool-aged children are generally consistent with the predictions of the implicature-based approach, while they require auxiliary assumptions on the ambiguity-based approach.

The rest of this chapter is organized as follows. In §2, we briefly outline the implicature approach and its predictions, and review the existing experimental work that has investigated the implicature approach. In §3, we briefly consider two implementations of the alternative, ambiguity-based approach and the predictions that they make. In §4, we discuss how the two theoretical approaches contend with the experimental findings to date, in particular focusing on the findings related to monotonicity, suspension and context dependence, and acquisition. §5 concludes with an outlook for future research.

2 The implicature approach

2.1 The implicature approach and its predictions

We have already seen that a sentence like (7a), repeated from above, gives rise to a multiplicity inference that Emily fed more than one giraffe. We can alternatively refer to the strengthened meaning of (7a) as that in (7b), once the multiplicity inference has been factored in.

(7)   a.   Emily fed giraffes.
        b.   Emily fed more than one giraffe.

We have also seen that this strengthened meaning disappears under negation, as (8a) does not seem to convey (8b). This behavior extends beyond the distinction between affirmative and negative declarative sentences to include questions (9) and antecedents of conditionals (10), suggesting sensitivity to a more general distinction between upward-entailing and downward-entailing contexts (though we will discuss some apparent exceptions in §3.2).

(8)   a.   Emily didn't feed giraffes.
        b.   Emily didn't feed more than one giraffe.

(9)   Did Emily feed giraffes?

(10)   If Emily fed giraffes, she will get a prize.

The implicature approach to the data above starts from the observation that this pattern is not unique to plurals, and is in fact parallel to the pattern of scalar implicatures more generally. For instance, a disjunctive sentence like (11a) gives rise to the strengthened exclusive meaning in (11b), but the latter disappears under negation, as shown in (12a-b); under the reading in (12b), (12a) would be compatible with Emily feeding both an elephant and a giraffe.

(11)   a.   Emily fed a giraffe or an elephant.
        b.   Emily didn't feed both a giraffe and an elephant.

(12)   a.   Emily didn't feed a giraffe or an elephant.
        b.   It's not the case that Emily fed a giraffe or an elephant but not both.

In response to these data, the most prominent proposal in the literature is to treat the multiplicity inference as a form of implicature. Under this approach, multiplicity inferences are not part of the literal meaning of plural morphology, but arise from the comparison of (13a) with (13b). That is, (13a) simply means that Emily fed one or more giraffes, and it is through the negation of the stronger (13b) that we derive the 'more than one' meaning.

(13)　　a.　Emily fed giraffes.
　　　　b.　Emily fed exactly one giraffe.

When (13b) is placed under negation, however, as in (14b), its meaning becomes weaker relative to the assertion in (14a), and therefore the implicature is not derived.

(14)　　a.　Emily didn't feed giraffes.
　　　　b.　It's not the case that Emily fed exactly one giraffe.

The story above can be implemented in various ways, and existing proposals in the literature range from treating the multiplicity meaning as a conversational implicature, arising from a Maximize Presupposition-based comparison with the presuppositional singular (something like (13b)) (Sauerland 2003, 2005), to treating it as a recursive scalar implicature (Spector 2007, Magri 2014) or an embedded scalar implicature (Zweig 2009, Ivlieva 2013, Mayr 2015). All of these different implementations of the implicature approach make the prediction in (15).

(15)　　**Uniformity Prediction:** If multiplicity inferences are implicatures, then everything being equal, multiplicity inferences and implicatures should exhibit similar behavioral profiles.

Now, given this prediction, it is important to note that adults' rates of implicature computation have been reported to vary quite widely across different lexical scales (van Tiel, van Miltenburg, Zevakhina & Geurts 2014). For this reason, some of the studies that we will turn to next involve comparisons across *different populations*. For instance, even if adults do not compute multiplicity inferences and classical scalar implicatures at precisely the same rates, we might expect that children initially exhibit similar performance on the two kinds of inferences, or at least that they will differ from adults in the same manner for the two kinds of inferences. By comparing across two populations, these studies are informative about the relationship between the two inferences, even if adults do not compute them at exactly the same rates.

2.2　Experimental work

We will now turn to a series of experiments that have tested the Uniformity Prediction in (15), both with adults and children.

2.2.1　Sauerland, Andersen, & Yatsushiro (2005)

Sauerland, Andersen & Yatsushiro (2005) present one of the first studies comparing adults' and children's performance on multiplicity inferences, with the aim of testing whether the latter can be analyzed as a form of implicature. If multiplicity

inferences are derived as implicatures, one might expect adults and children to treat the two kinds of inferences in a parallel fashion. While there is considerable variability in the developmental results concerning children's ability to compute scalar implicatures, a fairly consistent finding has been that children compute fewer classical scalar implicatures of the kind in (16a,c) than adults do (Gualmini, Crain, Meroni, Chierchia & Guasti 2001, Chierchia, Crain, Guasti & Thornton 2001, Noveck 2001, Papafragou & Musolino 2003, among many others). That is, children have been shown to accept sentences such as (16a) in situations where the stronger alternative (16b) is also true. This observation has usually been taken to indicate that children have failed to compute the implicature in (16c).

(16) a. Emily fed some of the giraffes.
   b. Emily fed all of the giraffes.
   c. Emily fed some but not all of the giraffes.

There has since been research indicating children can succeed on certain kinds of implicatures, in certain experimental contexts (see, for example, Miller, Schmitt, Chang & Munn 2005, Barner, Brooks & Bale 2011, Katsos & Bishop 2011, Tieu, Romoli, Zhou & Crain 2016 and Stiller, Goodman & Frank 2015). There is currently debate in the developmental literature as to the underlying source of children's apparently selective success across different implicatures; there is also disagreement as to whether non-implicature responses in children reflect a true absence of implicature. These debates are not necessarily important for the comparison we will make; we will simply rely on the generalization that without facilitatory experimental manipulations or specially selected scales, children behave systematically unlike adults in many apparently scalar implicature-based tasks. If this is so, and if we further assume that multiplicity inferences are implicatures, we are led to the Uniformity Prediction that children should also exhibit difficulty computing multiplicity inferences. In particular, they should compute fewer multiplicity inferences than adults, under the same kinds of conditions in which they compute fewer standard scalar implicatures than adults.

Sauerland et al. (2005) designed an experiment that involved a puppet playing the role of an alien asking questions about life on earth. In particular, the puppet would ask participants polar questions such as (17). The expectation was that a reading involving the multiplicity inference (e.g., *Does a dog have more than one tail?*) would lead to a negative answer, given common knowledge, while a literal interpretation of the question (e.g., *Does a dog have one or more tails?*) would lead to an affirmative response.

(17) Does a dog have tails?

The authors compared (17) to control questions like (18), in which the literal meaning of the plural is incompatible with world knowledge, and with (19), in which both the literal meaning of the plural and the multiplicity inference are compatible with world knowledge.

(18)   Does a boy have beaks?

(19)   Does a cat have feet?

The authors tested 14 children aged 3-5 years, and found the following pattern. Children were adult-like on the control questions, responding negatively to (18) and affirmatively to (19). However, they produced more affirmative responses to target questions like (17) than adults did, answering 'yes' to such questions 96% of the time. This would indicate that children accessed the literal meaning of the question in (17), i.e. *Does a dog have one or more tails?*

The observed difference between children and adults mirrors exactly the differences that have been reported on previous studies of implicatures, with children tending to accept implicature targets more than adults. Sauerland et al.'s (2005) results therefore provide a first piece of evidence in favor of the implicature approach to plurality.

However, as Sauerland et al. (2005) point out (see also Pearson et al. 2011, Tieu et al. 2014), their study has some potential limitations. First, multiplicity inferences are not always computed in polar questions, even for adults. Sauerland et al. (2005) discuss examples like (20), drawn from **?**, in which a response assuming a multiplicity inference from (20a) seems to be infelicitous.

(20)   a.   Does your office have windows?
        b.   Yes, (only one though).
        c.   #No, only one.

Sauerland et al. (2005) attribute the difference between a question like (20a) and one like their experimental (17) to whether the person asking the question already knows the answer to the question, with the relevant distinction between *true information-seeking questions* and *exam-type questions*. According to the authors, only *exam-type questions* systematically give rise to multiplicity inferences. As Pearson et al. (2011) point out, however, the experimental setting in Sauerland et al. (2005) was precisely one in which the puppet was asking true information-seeking questions; the puppet was ignorant about life on earth and wanted to acquire information by asking the participants questions about life on earth. It is therefore unclear that this distinction is relevant for explaining how the participants behaved in Sauerland et al.'s (2005) study.

Moreover, Sauerland et al.'s (2005) experimental stimuli involved generic in-

terpretations, which could have been misinterpreted by participants as containing dependent plurals. For example, children might have interpreted (17) along the lines of, *Do dogs have tails?*, in which case *yes*-responses would be entirely reasonable, independently of the children's ability to compute multiplicity inferences.

While the findings of Sauerland et al.'s (2005) study can be taken to provide some initial evidence in favor of an implicature approach, the results are not entirely conclusive for the reasons just discussed.

2.2.2 Pearson, Khan, & Snedeker (2011)

Pearson, Khan & Snedeker (2011) tackle the multiplicity-as-implicature hypothesis from a different angle, designing an experiment aimed to test whether such inferences are cancellable in upward-entailing contexts. The authors used a Covered Box task, modeled on Huang, Spelke & Snedeker (2013). Huang et al. (2013) presented participants with three boxes; the contents of two of the boxes were clearly visible to both the experimenter and the participant, while the contents of the third box were hidden. On each trial, participants were asked to give the experimenter the box that matched the experimenter's description. As an example, on one target trial, participants would see one box containing one cookie, one box containing three cookies, and a covered box. The experimenter would direct participants to "Give me the box with two cookies." The expectation was that if the exactly-*n* meaning of the numeral was a cancellable implicature, participants would select the box containing three cookies, thereby satisfying the experimenter's request. If, on the other hand, the exact meaning was not a cancellable implicature, but rather part of the semantics of the numeral, participants would necessarily have to select the covered box. Huang et al. (2013) found that adults chose the covered box on the numeral target trials, suggesting, at least *prima facie*, an exact semantics for the numeral. In contrast to the numeral trials, Huang et al. found that adults were more willing to cancel the scalar implicature of "some", for example selecting the box where Cookie Monster had *all* of the cookies when asked to give the box where Cookie Monster had "some of the cookies."

If the multiplicity inference is an implicature, just like the one arising from the scalar quantifier *some*, one should expect that in exactly the same experimental set-up, adults would also cancel the multiplicity inference in order to give the experimenter one of the visible boxes, just as was observed for *some* in Huang et al.'s (2013) study. This is precisely what Pearson et al. (2011) set out to investigate. They presented participants with three cards at a time, two visible and one face down. On one critical target, for instance, the two visible cards contained Big Bird with no kite and Big Bird with just one kite, respectively. Participants then heard the instruction in (21).

(21)   Point to the card where Big Bird has kites.

The expectation was that if participants could cancel the multiplicity inference, interpreting (21) as *Point to the card where Big Bird has one or more kites*, they would be happy to select the visible picture where Big Bird had only one kite. On the other hand, if participants interpreted (21) with the multiplicity inference (i.e. *Point to the card where Big Bird has more than one kite*), they would have to select the covered picture.

Pearson et al.'s (2011) study consisted of two experiments. In the first experiment, 16 native speakers of English were presented with sentences like (21) in target conditions like the one described above. Participants displayed an overwhelming preference for the face down card on target trials, in parallel with Huang et al.'s (2013) results for the numeral targets but not the *some* targets. *Prima facie*, such a finding goes against the implicature approach, which would predict that the multiplicity inference should be cancellable in the same contexts in which the *some-not-all* implicature is cancelled.

In a follow-up experiment, Pearson et al. (2011) introduced the term "only" into the test sentences, as in (22), and found that participants were much happier to select the visible picture in which Big Bird had exactly one kite.

(22)   Point to the card where Big Bird only has kites.

The authors speculate that this result may be due to the fact that on a standard semantics of "only," the multiplicity inference is merely presupposed, and not asserted. Even if this is so, we are nevertheless left with a difference that is puzzling on the implicature approach: when presented in the same experimental set-up, the implicature of *some*, but not the multiplicity inference, is easily cancellable – even in the absence of "only."

In sum, Pearson et al.'s (2011) results may be suggestive of the implicature approach, but do not argue conclusively in favor of the account, due to the observed difference between multiplicity inferences and scalar implicatures with respect to ease of cancellability.

2.2.3   Tieu, Bill, Romoli, & Crain (2014, 2015, 2017)

Following up on Sauerland et al. (2005), Tieu, Bill, Romoli & Crain (2014, 2015, Under review) conducted another pair of acquisition experiments aimed at detecting multiplicity inferences in preschool-aged children. These authors set out to test the developmental uniformity prediction that arises from the scalar implicature approach to multiplicity inferences, namely that the pattern of children's multiplicity inferences should mirror their performance on standard scalar implicatures. On the

implicature account, any observed differences in the acquisition profiles of the two inferences would minimally have to be explained via auxiliary assumptions.

In order to address some of the potential limitations of the Sauerland et al. (2005) study, Tieu et al. (2014) avoided the use of polar questions and instead tested the interpretation of the plural in affirmative and negative declarative sentences. By manipulating monotonicity, the authors could further test the implicature approach advanced by Spector (2007), which would predict the disappearance of multiplicity inferences in downward-entailing environments (an exception to this pattern could arise through the presumably less preferred computation of the multiplicity inference *in the scope of negation*, a point we return to below).

The authors used a Truth Value Judgment Task (Crain & Thornton 1998), in which participants watched a series of short stories and then had to judge a puppet's descriptions of the stories. A sample target story is provided in (23); in this story, the protagonist fed exactly one pig. Participants were expected to accept the accompanying upward-entailing target sentence (23a) if they accessed the literal existential meaning, but they were expected to reject the sentence if they accessed the multiplicity inference. The downward-entailing counterpart of (23a) can be found in (23b) (though the stories were changed between the upward- and downward-entailing conditions, to keep the task interesting for the child participants). Participants were expected to reject the downward-entailing target in (23b) if they accessed the literal meaning of the test sentence, since it was false that Emily didn't feed any pigs (she fed one). They could, however, accept the sentence if they computed the multiplicity inference locally under negation, giving rise to the meaning *Emily didn't feed more than one pig*.

(23)   *Story:* Emily is visiting the pig farm today. It's lunchtime for the pigs. Emily has an apple, and that's just enough to feed the first pig! Oh no! What about the other pigs? The farmer says, "That's okay, Emily! I'll feed the others later!" So in the end, Emily only fed this pig!
EXPERIMENTER: Hey Ellie, what happened in the story?

    a.   PUPPET: Emily fed pigs!   *Plural UE target*
    b.   PUPPET: Emily didn't feed pigs!   *Plural DE target*

In their first experiment, Tieu et al. (2014) tested 28 English-speaking 4- and 5-year-olds and 43 adult native speakers of English. The adult participants rejected the upward-entailing targets 92% of the time, but accepted the downward-entailing targets 42% of the time. Adults, then, computed more multiplicity inferences in upward-entailing than in downward-entailing environments, but also computed multiplicity inferences under negation to some degree (a less plausible source of the *yes*-responses to the negative sentences would be a wide scope reading of the

bare plural above negation, although bare plurals are generally assumed to resist wide scope, cf. Carlson 1977). Children, on the other hand, computed multiplicity inferences only 40% of the time in the upward-entailing condition, and 19% of the time in the downward-entailing condition.

In a second experiment, Tieu et al. (2015) conducted a more direct comparison of the plurality inference and classical scalar implicatures by including a *some* implicature condition. The plurality inference condition was similar in structure to the that of the first experiment, with the materials modified slightly to ensure maximal parallelism between the plural and *some* condition. A plural target can be found in (24) and a scalar implicature target in (25). In both cases, participants were expected to accept the sentence if they accessed the literal meaning, and to reject the sentence if they computed the relevant inferences in (24b) and (25b).

(24)   *Scenario:* Zebra is at the orchard. There are four bananas in one tree and four oranges in another tree. Zebra picks one of the four bananas.

    a.   Zebra picked bananas.

    b.   ⤳ *Zebra picked more than one banana*

(25)   *Scenario:* Lion is at the orchard. There are four oranges on the ground and four apples on the ground. Lion carries four of the four apples.

    a.   Lion carried some of the apples.

    b.   ⤳ *Lion carried some but not all of the apples*

Tieu et al. (2015) tested 17 English-speaking 4- and 5-year-olds and 27 adults. Similarly to the first experiment, adults computed more multiplicity inferences in the upward- than in the downward-entailing environment (75% vs. 19%), and children computed far fewer multiplicity inferences than adults regardless of monotonicity (16% in the upward-entailing condition vs. 4% in the downward-entailing condition). The child participants also computed fewer implicatures from *some* than adults (adults 81% vs. children 28%), consistent with the pattern from multiplicity inferences. Children's responses to the two kinds of inference targets were also significantly correlated. On the whole, the results from the two experiments appear to be consistent with the implicature approach to plurality.

2.2.4   Patson (2016)

Patson (2016), building on Patson, George & Warren (2014), conducted two picture-matching task experiments aimed at probing the interpretations of sentences containing plural definites like "the leaves." The adult participants read sentences that implied either a particular spatial distribution of the objects denoted by the definite description (Experiment 1), or different set sizes (Experiment 2). For example, in

Experiment 1, "the leaves" could be implied to be spatially gathered, as in (26), spatially distributed, as in (27), or neutral with respect to spatial distribution (28):

(26)   The gardener raked up the leaves.

(27)   The breeze scattered the leaves.

(28)   Thomas was confused about the shape of the leaves.

After reading the test sentence, participants were presented with a picture of the denoted object, and they had to decide if the picture was of an object that had been mentioned in the sentence. Crucially, the picture either matched in spatial distribution, mismatched in spatial distribution, or was neutral with respect to spatial distribution (e.g., a pile of leaves; a group of scattered leaves; a single leaf).

Patson observed high accuracy in the picture judgment task, with participants matching the 'collective' pictures with the 'collective' sentences, and the spatially distributed pictures with the 'scattered' sentences. Patson also collected reaction times for participants' responses, and found that participants were quicker to accept pictures that matched the sentences in terms of spatial distribution than pictures that mismatched the sentences. Crucially, participants were as fast to accept the *matching* pictures as they were to accept the *singular* pictures, suggesting the participants had activated the 'at least one' interpretation of the plural definite descriptions.

In the second experiment, Patson manipulated the implied set size of the objects rather than spatial distribution. The test sentences either implied a small or large set size (e.g., *The scientist examined the tray of leaves* vs. *During autumn the yard is full of leaves*) or was neutral with respect to set size (e.g., *I like the smell of leaves*). The pictures accordingly depicted, e.g., a small set of leaves, a large (uncountable) set of leaves, or a single leaf. Patson again observed high accuracy rates across conditions. Crucially, the reaction time results replicated those of the first experiment: participants were as fast to respond to pictures that matched the implied set size as they were to accept the singular picture.

These results replicate and extend those in Patson et al. (2014). The findings suggest that when participants represent plural definite descriptions they also activate the corresponding singular meaning. As Patson argues, such findings are in line with the hypothesis that the plural is number-neutral in its literal meaning, and is associated with a multiplicity inference via an implicature. By contrast, the results are much harder to reconcile with a semantic proposal on which the multiplicity meaning is encoded in the literal meaning of the plural definite description. Had the plural definite descriptions only activated representations of multiple objects, one might have expected participants to be slower to respond to the singular pictures than to the pictures containing multiple objects.

3 Ambiguity-based approaches and their predictions

In addition to approaches that treat the 'more than one' meaning of plural morphology as an *implicature*, there exist ambiguity approaches that essentially posit two possible meanings for the plural, with some additional principle for selecting between two meanings. Compared to the implicature approach, these approaches have been relatively less investigated from an experimental perspective. Nevertheless, in the following subsections we present an overview of the existing theoretical approaches and their accompanying experimental investigations where applicable (this discussion can also be found in detail in Tieu et al. Under review).

3.1 Farkas & de Swart (2010)

Farkas & de Swart (2010) propose that plural morphology is polysemous between a weak *inclusive* meaning and a strong *exclusive* meaning. A plural sentence like (29), repeated from above, can thus be associated with two possible readings: a weak reading along the lines of (30a) and a strong reading along the lines of (30b).

(29)   Emily fed giraffes.

(30)   a.   Emily fed one or more giraffes.   WEAK
         b.   Emily fed more than one giraffe.   STRONG

Under negation, the relative strength of the two readings is reversed; the negative (31) can be associated with the strong (32a) or the weak (32b).

(31)   Emily didn't feed giraffes.

(32)   a.   Emily didn't feed one or more giraffes.   STRONG
         b.   Emily didn't feed more than one giraffe.   WEAK

As we have seen from our previous discussion, people have rather strong tendencies to interpret (29) and (31) in a particular way, typically favoring (30b) and (32a), respectively. To explain these interpretive preferences, Farkas & de Swart (2010) assume the Strongest Meaning Hypothesis in (33) (see Dalrymple, Kanazawa, Kim, Mchombo & Peters 1998, Winter 2001, among others).

(33)   *The Strongest Meaning Hypothesis for Plurals*: for a sentence involving a plural nominal, prefer that interpretation of plural which leads to the stronger overall interpretation for the sentence as a whole, unless this interpretation conflicts with the context of utterance.

With (33), the ambiguity approach can capture the monotonicity effect we observe: the principle favors the interpretation (30b) for the positive (29) and (32a) for the

negative (31), as these correspond to the strongest possible interpretations of the respective sentences.

Note that the principle in (33) will also allow for the weaker interpretations to emerge if the stronger ones are in conflict with the context of utterance. For instance, a weak interpretation can be forced by adding a continuation that is in conflict with the strong reading, thereby blocking it, as illustrated in (34); here the negated plural receives the weak interpretation along the lines of (32b), rather than the strong interpretation in (32a).

(34)   Emily didn't feed giraffe**s**, because she fed only one!

Like the implicature approach, the ambiguity approach can therefore capture the monotonicity effect and the presence of readings like that in (34). That is, the two approaches make the same predictions when it comes to the interpretation of plural morphology in positive and negative sentences, and they likewise make the same prediction regarding the availability of the reading in (34).

One place where the two theories do make different predictions is in the relationship between multiplicity and implicatures. Unlike the implicature approach, the ambiguity approach makes no particular predictions regarding this relationship, since the two phenomena are not seen as related. This relationship between multiplicity and implicatures therefore in principle provides a fruitful avenue for potentially adjudicating between the two approaches. One complication to note, however, is the observation that there is considerable variation in implicature computation rates across scales (van Tiel et al. 2014); even if the multiplicity meaning arises as an implicature, we might not expect to observe it occurring at similar rates as standard cases of implicature.

A further avenue for investigating the predictions of the theories lies in child language acquisition. As Tieu et al. (Under review) discuss, the ambiguity approach makes certain predictions about how plurality should be acquired. On this approach, children must not only acquire the two proposed meanings of the plural, they must also be able to make use of the Strongest Meaning principle in an adult-like manner, in order to resolve the ambiguity. Tieu et al. (Under review) present three possible learning stages during which children might be observed to be non-adult-like in their interpretation of the plural. First, they might only have the strong meaning of the plural, in which case we should expect them to be adult-like on the plural in upward-entailing environments but not in downward-entailing environments. Second, they might only have the weak meaning of the plural, in which case they should be adult-like only in downward-entailing environments. Finally, they might have both meanings for the plural, but cannot yet deploy the Strongest Meaning principle as adults do. They might instead, for example, be guided by charity, selecting the

interpretation that is made true in the context.

Summarizing, the ambiguity approach of Farkas & de Swart (2010) and the implicature approach both predict more multiplicity inferences in upward-entailing than in downward-entailing environments, but they differ in how they view the relationship between multiplicity and implicatures. Furthermore, they differ in the predictions that they make for how children and adults should perform on multiplicity inferences (and scalar implicatures). In the next subsection, we turn to a different incarnation of the ambiguity approach, based not on monotonicity but rather on the distinction between particularized and general reference.

3.2 Grimm (2013)

While many of the pre-existing accounts of the plural meaning place importance on monotonicity, Grimm (2013) presents an ambiguity analysis based on the distinction between particularized and general reference. Grimm points out that the inclusive plural meaning can arise in environments like (35), which can be shown to be neither strictly downward-entailing (36), nor to block scalar implicatures (37).

(35)   Sherlock Holmes should question local residents to find the thief.
        ↛ Sherlock Holmes should question groups of two local residents or more

(36)   Sherlock Holmes should question local residents.
        ↛ Sherlock Holmes should question local residents with red hair

(37)   Sherlock Holmes should question three local residents.
        ⤳ Sherlock Holmes should question exactly three local residents

Note that the fact that the multiplicity inference appears to be context-dependent is not *per se* a challenge to the implicature approach. Indeed, the latter can readily account for context-dependence by appealing to relevance. That is, the data above are challenging only if it can be shown that the *exactly-one* alternative is relevant in the context against which (35) is evaluated.

Grimm takes a different approach altogether and argues that, rather than downward-entailingness, the relevant property to account for the distribution of the inclusive plural reading is the weakened referential status of the nominal. That is, the inclusive plural reading arises in contexts where the nominal doesn't presuppose the existence of any particular referents. Specifically, Grimm considers the relevant distinction to be between *object-level* and *concept-level* entities.

Grimm sets out to experimentally probe the inclusive meaning of plurals in interrogative contexts, by conducting two experiments involving plural nouns in two kinds of contexts. In the first experiment, participants were presented with pictures which clearly instantiated the object under question. For example, participants would

see a photo of a woman holding either one mug or two mugs, and would have to answer the *yes/no* question *Is the woman in this picture holding mugs?* In the second experiment, participants were presented with "rules-and-regulations" contexts meant to foster generic or kind-level construals (Carlson 1995). For example, participants saw questions meant to evoke a corporate environment, such as (38).

(38) *Did your team terminate projects this fiscal quarter?*
EMPLOYEE FACTS: employee's team has terminated exactly one project this fiscal quarter
What answer should the employee give? *Yes/No*

Grimm reports that in the first experiment, participants provided *yes*-answers 92% of the time when the picture contained multiple objects, but gave *yes*-answers only 32% of the time when presented with a single object. In contrast, in the second experiment, adults answered *yes* 99% of the time when the "employee facts" line contained a plural entity like *three projects*, and crucially still preferred the inclusive interpretation on the critical targets (38), responding *yes* 78% of the time.

Grimm's experimental data reveal that participants disfavor inclusive readings of the plural when considering object-referring uses of plural nouns, but favor inclusive readings in contexts that encourage general reference. To capture the data, Grimm proposes an analysis that integrates Krifka's (1995) analysis of common nouns and Roberts's (1996) framework for discourse structure involving the Question Under Discussion (QUD).

First, Grimm draws on Krifka (1995) to provide the two representations needed for the two possible interpretations for a given plural noun: a *quantified object* reading and an *instance of a concept* reading (39b).

(39) a. Quantified Object: $\lambda n \lambda i \lambda x [R_i(x, DOG) \wedge OU_i(DOG)(x) = n]$
      b. Instances of a Concept: $\lambda i \lambda x \exists n [R_i(x, DOG) \wedge OU_i(DOG)(x) = n]$

These two representations then provide two possible interpretations for a given plural noun. The plural marker can be interpreted as a number marker (which Grimm diverges from Krifka in assuming is interpreted as *more than one*), as in (40a); in this case the interpretation is true when the noun's referent is the type of thing named and has cardinality greater than one. Alternatively, the plural noun can enter the derivation as a kind or concept, whereupon the interpretation succeeds when at least one entity realizes the kind *dog*, as in (40b); in this case, the reading is number-neutral, being true as soon as the noun's referent is the type of thing named.

(40) a. $[\![\text{dogs}]\!] := \lambda i \lambda x [R_i(x, DOG) \wedge OU_i(DOG)(x) \geq 2]$
      b. $[\![\text{dogs}]\!] := \lambda i \lambda x \exists n [R_i(x, DOG) \wedge OU_i(DOG)(x) = n]$

15

Next, Grimm situates his analysis within the discourse framework of Roberts (1996). In essence, distinct QUDs can be derived that query either the existence or the number of objects, respectively:

(41)　a.　Did Ed see dogs?　　　　　　　　　　　　　　　(Existential QUD)
　　　　　　$?(\lambda i \exists x \exists n \,[\text{Ed see } x \text{ in } i \wedge R_i(x, \text{DOG}) \wedge OU_i(\text{DOG})(x) = n])$
　　　　b.　Did Ed see dogs?　　　　　　　　　　　　　(Quantitative QUD)
　　　　　　$?(\lambda i \exists x \,[\text{Ed see } x \text{ in } i \wedge R_i(x, \text{DOG}) \wedge OU_i(\text{DOG})(x) \geq 2])$

According to Grimm, the interpretation of the plural noun is then guided by which of (41a,b) corresponds to the immediate QUD that the conversational participants are attempting to answer. When reference to particular entities is not yet determined, the Existential QUD is appropriate, licensing an inclusive reading; when reference to particular entities is determined, it is the Quantitative QUD that is the immediate QUD, and the exclusive reading is licensed. This would explain why in Grimm's first experiment, where participants were presented with particular objects (i.e. one or more mugs), participants preferred the Quantitative QUD and therefore the exclusive reading of the plural, whereas in the second experiment, where there was no presumption that any projects were completed, participants preferred the Existential QUD and therefore the inclusive reading of the plural.

4　General discussion

The existing experimental investigations into the semantics and pragmatics of plurality yield a number of insights, all the while raising new questions for further research. We now turn to these insights and remaining puzzles.

4.1　Sensitivity to monotonicity

The multiplicity inference appears to be sensitive to monotonicity, as most clearly evidenced by the data in Tieu et al. (2014, 2015). The adult participants in these experiments generally preferred to reject both positive and negative plural descriptions like (42a,b) in singular contexts, i.e. in which Emily fed exactly one giraffe. Such responses indicate that adults interpreted the plural exclusively in the upward-entailing (42a), but inclusively in the downward-entailing (42b).

(42)　a.　Emily fed giraffes.
　　　　b.　Emily didn't feed giraffes.

As we have seen, both the implicature approach and the ambiguity approach of Farkas & de Swart (2010) predict precisely this pattern, the former through implicatures

that arise in upward-entailing but not downward-entailing environments, and the latter through a Strongest Meaning principle. Now consider the ambiguity approach of Grimm (2013). On Grimm's proposal, participants' interpretation of the plural nouns would have been guided by whether they accessed the Existential QUD or the Quantitative QUD. However, Tieu et al. (2015) used the same kinds of stories for their upward-entailing and downward-entailing targets. Given the stories were the same, presumably the salient QUDs were also equivalent across the two conditions. It is therefore somewhat puzzling that adults, particularly in the second experiment, computed far fewer multiplicity inferences under negation (75% vs. 19%). One would have to say that the positive and negative sentences themselves invoked different QUDs, and therefore different interpretations of the plural, since only the sentences differed between conditions (REFERENCE NEGATION CHAPTER).

Future research might further investigate the respective roles that the QUD and monotonicity may play in adults' interpretation of plural morphology, and the interaction between the two factors (if any). For the moment, it is clear that adults distinguish between upward- and downward-entailing contexts when interpreting plural morphology. This sensitivity to polarity is expected under all of the various incarnations of the implicature approach in the literature and on certain implementations of the ambiguity approach.

4.2   Suspension and context dependence

The theoretical approaches we have discussed are not quite so straightforward to reconcile with the experimental evidence pertaining to suspension and cancellability. First, we have discussed the observation that multiplicity inferences are harder to suspend than regular scalar implicatures, as can be seen by the contrast between (43) and (44).

(43)   Some of the students came. In fact, all of them did.

(44)   Emily fed giraffes. #In fact, she fed exactly one.

This difference is reflected in Pearson et al.'s (2011) results, especially those of their first experiment. While their second experiment reveals that multiplicity inferences are easier to suspend in the scope of a presupposition trigger, the exact nature of the interaction between presuppositions and implicatures in such cases is not very clear, rendering it difficult to draw conclusions regarding suspension of multiplicity inferences from this particular data point.

The challenge for the implicature approach is to account for the difficulty of suspending multiplicity inferences in cases like (44) above, while at the same time accounting for their apparent context dependency in cases like Grimm's (35). More

generally, it remains an open question how to account for the observed difference in cancellability between scalar implicatures of terms like *some* and multiplicity inferences, as well as the role played by presuppositions in apparent cases of suspension.

As Pearson et al. (2011) discuss, the observation that multiplicity inferences are difficult to cancel has already been made in the literature. Chierchia, Fox & Spector (2011), in particular, observe the oddness of a case like (44), and propose that the implicature triggered by (44) is in fact *obligatory* (Chierchia et al. 2011). That is, the implicature arising from *some* in (43) is context-dependent and easily suspended, while the multiplicity inference arising from the plural in (44) is not. An outstanding question is how exactly to reconcile this with Grimm's (2013) data, which suggest context dependency of the multiplicity inference.

As mentioned, Grimm's (2013) observation that multiplicity inferences are context-dependent is not in and of itself a problem for the implicature approach. The multiplicity inference could be treated as an optional implicature, the presence of which is dependent on the context. In particular, the presence of the implicature would depend on whether the *exactly-one* alternative is made relevant in the context. In Grimm's first experiment, this alternative was made relevant in the context, while in the second experiment it was not. In this way, the implicature approach can explain both the sensitivity to polarity and the context dependency displayed by the multiplicity inference. One would, however, have to work through the details of how to reconcile cases where context dependence is evident (like Grimm's (35)), with the apparent lack of suspension in cases like (44). Notice that the same issue arises on the ambiguity approach of Farkas & de Swart (2010), which would need to explain why the Strongest Meaning principle appears to be violable in cases like (35) but not in cases like (44).

4.3   Children's development of plural meanings

The data pertaining to monotonicity, suspension, and context dependence do not unequivocally argue for one approach to multiplicity over the other. However, considerations of language acquisition and learnability may be able to shed light on the debate. Let us consider how each approach might explain the experimental results we have discussed.

We have observed that children do not compute as many multiplicity inferences as adults do (e.g., Sauerland et al. 2005, Tieu et al. 2014). What could be missing, according to the implicature approach? On the implicature analysis, the child must learn the relevant alternatives, contrasting the plural form with some version of the singular; otherwise no inference can arise from the plural. Children who cannot yet compute multiplicity inferences may lack knowledge of or access to the required alternatives. In fact, access to alternatives has been proposed to be a problem for

children's generation of implicatures more generally; in cases where children *do* have access to the required alternatives, research indicates they are fully capable of generating implicatures and implicature-like meanings (see, for example, Barner, Brooks & Bale 2011, Stiller, Goodman & Frank 2015, Tieu, Romoli, Zhou & Crain 2016 and Singh, Wexler, Astle-Rahim, Kamawar & Fox 2016). On the implicature approach, the multiplicity inference may be particularly challenging for the child learner of English, given that neither the plural morpheme *-s* nor the (enriched) singular alternative correspond to standalone lexical items that the child can easily contrast with each other (unlike the cases of, for example, "some"/"all", "or"/"and", "might"/"must", and numerals).

How might the ambiguity approach of Farkas & de Swart (2010) contend with the developmental data? On the whole, the child data appear to pose a challenge to this approach. While the ambiguity approach is not incompatible with similar developmental profiles for multiplicity inferences and standard implicatures, it would not appear to readily account for children's behavior across the polarities, that is, in positive and negative sentences. Recall that on the ambiguity approach, the positive (45a) is ambiguous between the readings in (45b) and (45c), while the negative (46a) is ambiguous between the readings in (46b) and (46c). The Strongest Meaning hypothesis is expected to favor (45b) for (45a) and (46c) for (46a).

(45)   a.   Emily fed giraffes.
           b.   Emily fed more than one giraffe.   STRONG
           c.   Emily fed one or more giraffes.   WEAK

(46)   a.   Emily didn't feed giraffes.
           b.   Emily didn't feed more than one giraffe.   WEAK
           c.   Emily didn't feed one or more giraffes.   STRONG

As we discussed in §3.1 (following discussion in Tieu et al. Under review), there are three main developmental scenarios that might give rise to non-adult-like behavior on the ambiguity approach. Children might access only the weak meaning of the plural; they might access only the strong meaning of the plural; or they might have both meanings but fail to apply the Strongest Meaning hypothesis in an adult-like way. Yet what the child data in Tieu et al. (2014, 2015) reveal is that children in fact favor the weak reading of the positive (45a), but the strong reading of the negative (46a). It is unclear how the ambiguity approach might capture this observed pattern across the two polarities. As suggested by Tieu et al. (Under review), for example, the ambiguity approach might posit that children engage with the Strongest Meaning principle differently from adults, obeying the principle only in downward-entailing contexts. But while it is plausible that children at a certain age might differ from adults in the deployment of a pragmatic principle like the Strongest

Meaning hypothesis, it's not clear why their application of such a principle would vary systematically with the polarity of the context.

Alternatively, the ambiguity approach might explain the data by positing that children in this age range have only mastered the weak *inclusive* meaning of plural morphology, hence accessing a globally *weak* reading of the positive (45a) but a globally *strong* reading of the negative (46a). But as Tieu et al. (Under review) discuss, such a scenario raises a potential learnability problem, in the form of a subset problem (Berwick 1985, Crain, Ni & Conway 1994, Gualmini & Schwarz 2009). In particular, if a child has acquired the weak meaning of the plural morpheme, any evidence they encounter for the strong meaning will also be compatible with the weak meaning, given that the strong meaning entails the weak meaning. What evidence would then prompt the child to move from a weak to a strong meaning for the plural morpheme?

Finally, let us consider what the child has to learn on an account like Grimm's. One possible explanation for children's difficulty with multiplicity inferences, under Grimm's proposal, might lie in a difficulty with selecting the correct QUD, and thereby accessing the target meaning of the plural. In this scenario, what the learner needs to develop is something more general than a particular kind of inference; the child has yet to become adult-like in her management of the discourse structure (see Zondervan 2007, 2009, 2010 for discussion on the role of QUD in implicature computation, and Gualmini, Hulsey, Hacquard & Fox 2008, Conroy, Takahashi, Lidz & Phillips 2009, Lewis 2013 for relevant discussion of the role of QUD in child language experiments). In the Tieu et al. (2014, 2015) experiments, for instance, one might posit that children favored the Existential QUD, while adults favored the Quantitative QUD; what would develop over time is an alignment with the adult-like selection of the target QUD.[2]

Alternatively, as on the ambiguity approach of Farkas & de Swart (2010), Grimm's approach might also posit that children simply haven't yet mastered both meanings or representations of the plural. Given their preference for the inclusive meaning of the plural, one might posit that children acquire the kind reading earlier than the quantified object reading (see Sneed 2005, 2008, Lazaridou-Chatzigoga, Katsos & Stockall 2015 for discussion of children's acquisition of generic noun phrases). Future work might more closely investigate the developmental connection between plurality and genericity.

---

2 It is worth pointing out, however, that the critical stories in the Tieu et al. experiments were such that the protagonist engaged with a particular entity that was explicitly mentioned; according to Grimm (2013), this therefore should have activated the Quantitative QUD as the immediate QUD. The justifications that children provided for accepting (the inclusive reading of) the plural in sentences like *Emily fed pigs* (where Emily fed exactly one pig), do not seem consistent with this (e.g., *Because she feed a pig / Because she said the pig has been feeded, and that happened*).

Summarizing, the existing experimental data from children appear to be more straightforwardly accounted for under the implicature approach. However, as we have discussed, both approaches to multiplicity inferences provide possible avenues for further investigating the developmental path that children may take to an adult-like understanding of plural morphology.

Before closing this section, it is worth noting that the acquisition studies by Sauerland et al. (2005) and Tieu et al. (2015) pertained specifically to preschoolers' knowledge of multiplicity inferences. This age range (3–5 years) is apt for investigating connections between multiplicity inferences and scalar implicatures, as previous studies of implicature have likewise targeted this age range, and shown that preschoolers typically compute fewer standard implicatures than adults do. On the face of it, the findings of Sauerland et al. (2005) and Tieu et al. (2015) appear to support the hypothesis that multiplicity inferences are a form of implicature, with similar developmental profiles. However, the observation that 4-year-old children may not be fully adult-like in their interpretation of plural morphology is *prima facie* at odds with the findings of studies conducted with much younger children. In particular, it has been reported that children begin producing plural morphemes at 22 months of age (e.g., Brown 1973, Mervis & Johnson 1991, Fenson, Dale, Reznick, Bates, Thal & Pethick 1994, Barner, Thalwitz, Wood & Carey 2007),[3] and that 24-month-olds are sensitive to plural marking in comprehension (Wood, Kouider & Carey 2009, Davies, Xu Rattanasone & Demuth To appear). For example, Wood, Kouider & Carey (2009) used a manual search paradigm in which 2-year-old children were prompted to search for objects in a box. The authors found that the 24-month-old participants searched longer in the plural condition, suggesting they were expecting to see 'more than one' object in such cases. More recent work by Davies, Xu Rattanasone & Demuth (To appear) also suggests that English-acquiring 24-month-olds have an understanding of plural morphology (in particular, of the voiceless /s/ allomorph).

Tieu et al. (Under review) attempt to explain the difference in the findings from 2-year-olds and 4-year-olds by identifying differences in the experimental paradigms used to test the two age groups. Studies like Wood et al. (2009) (and the more recent Davies et al. To appear) allow toddlers to derive a meaning for the linguistic description, and require the toddler to execute an action (or looks to a particular object) based on that interpretation. By contrast, the Truth Value Judgment Tasks in Tieu et al. (2014, 2015) presented 4-year-olds with situations in which only a single object was depicted. The acceptance of a plural description in such a situation doesn't

---

3 There are also reports, however, that children do not always produce the plural in required contexts (for discussion, see Berko 1958, Cazden 1968, Brown 1973, Mervis & Johnson 1991, Marcus, Pinker, Ullman, Hollander, Rosen & Xu 1992, Marchman, Plunkett & Goodman 1997, Winitz, Sanders & Kort 1981, Zapf & Smith 2008).

necessarily mean that children didn't *expect* there to be more than one object. The 4-year-olds may simply have been willing to override their expectation, and accept the use of plural morphology in a context in which they themselves would not have produced the plural. As Tieu et al. (Under review) point out, this acquisition scenario is quite consistent with a recent proposal regarding children's so-called *pragmatic tolerance* when it comes to scalar implicatures (Katsos & Bishop 2011). According to Katsos & Bishop (2011), preschoolers are simply more tolerant and charitable than adults when forced to accept or reject underinformative scalar descriptions.

5   Concluding remarks

In this chapter, we have presented an overview of recent experimental work investigating the semantics and pragmatics of plurality. These studies have examined the 'more than one' meaning that is seemingly contributed by plural morphology, referred to in the literature as a *plurality* or a *multiplicity inference*. As we have discussed, theoretical treatments of the multiplicity inference include deriving the 'more than one' meaning as a kind of implicature (Sauerland 2003, 2005, Spector 2007, Zweig 2009, Ivlieva 2013, Magri 2014, Mayr 2015), or treating the plural as ambiguous between two possible meanings (Farkas & de Swart 2010, Grimm 2013). Recent findings from experiments conducted with children and adults reveal that the plural meaning is sensitive to monotonicity, context-dependent, suspendable under certain conditions, and computed by preschool-aged children on a par with standard implicatures (Sauerland 2005, Pearson et al. 2011, Tieu et al. 2014, 2015, Grimm 2013). As we have discussed, the implicature approach to multiplicity inferences, in all of its different incarnations, is largely consistent with the experimental evidence accumulated in the literature in recent years, in particular that arising from the comparison between preschool-aged children and adults. This body of experimental work, however, has also introduced certain challenges, arising in particular from the way these inferences can be suspended, their apparent context dependency, and the relationship between multiplicity inferences and presuppositions. In addition, while some initial attempts at comparing multiplicity inferences and classical scalar implicatures suggest that the two exhibit similar behavioral profiles, more work is needed in order to understand the exact relationship between multiplicity inferences and standard cases of implicature, in both the developing and the adult grammar.

References

Barner, D., D. Thalwitz, J. Wood & S. Carey. 2007. On the relation between the acquisition of singular-plural morpho-syntax and the conceptual distinction between one and more than one. *Developmental Science* 10. 365–373.

Barner, David, Neon Brooks & Alan Bale. 2011. Accessing the unsaid: the role of scalar alternatives in children's pragmatic inferences. *Cognition* 118(1). 84–93.

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14. 150–177.

Berwick, Robert C. 1985. *The acquisition of syntactic knowledge*. Cambridge, Massachusetts: MIT Press.

Brown, Roger. 1973. *A first language: The early stages*. London: Allen and Unwin.

Carlson, Greg N. 1995. Truth-conditions of generic sentences. In Greg N. Carlson & Francis Jeffrey Pelletier (eds.), *The generic book*, 224–237. Chicago: Chicago University Press.

Carlson, Gregory N. 1977. *Reference to kinds in English*: University of Massachusetts, Amherst dissertation.

Cazden, Courtney B. 1968. The acquisition of noun and verb inflections. *Child Development* 39. 433–448.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6(4). 339–405.

Chierchia, Gennaro, Stephen Crain, Maria Teresa Guasti & Rosalind Thornton. 2001. 'Some' and 'or': A study on the emergence of logical form. In S. Catherine Howell, Sarah A. Fish & Thea Keith-Lucas (eds.), *Boston University Conference on Language Development (BUCLD) 24*, 22–44. Somerville, Massachusetts: Cascadilla Press.

Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2011. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *An international handbook of natural language meaning volume 3*, Berlin: Mouton de Gruyter.

Conroy, Anastasia, Eri Takahashi, Jeffrey Lidz & Colin Phillips. 2009. Equal treatment for all antecedents: How children succeed with Principle B. *Linguistic Inquiry* 40(3). 446–486.

Crain, Stephen, Weijia Ni & Laura Conway. 1994. Learning, parsing, and modularity. In Jr. Clifton, Charles, Lyn Frazier & Keith Rayner (eds.), *Perspectives on sentence processing*, 443–467. Hillsday, New Jersey: Lawrence Erlbaum.

Crain, Stephen & Rosalind Thornton. 1998. *Investigations in Universal Grammar: A guide to experiments on the acquisition of syntax and semantics*. Cambridge, Massachusetts: MIT Press.

Dalrymple, M., M. Kanazawa, Y. Kim, S. Mchombo & S. Peters. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistic and Philosophy* 21. 159–210.

Davies, Benjamin, Nan Xu Rattanasone & Katherine Demuth. To appear. Two-year-olds' sensitivity to inflectional plural morphology: Allomorphic effects. *Language Learning and Development* .

Farkas, Donka & Henriette de Swart. 2010. The semantics and pragmatics of plurals.

Semantics & Pragmatics 3(6). 1–54.

Fenson, L., P.S. Dale, J. Reznick, E. Bates, D.J. Thal & S.J. Pethick. 1994. Variability in early communicative development. *Monograph of the Society for Research in Child Development* 59(5, Serial No. 242).

Grimm, Scott. 2013. Plurality is distinct from number-neutrality. In *Proceedings of the north east linguistic society 41*, 247–258. Amherst, MA: University of Massachussetts at Amherst.

Gualmini, Andrea, Stephen Crain, Luisa Meroni, Gennaro Chierchia & Maria Teresa Guasti. 2001. At the semantics/pragmatics interface in child language. In Rachel Hastings, Brendan Jackson & Zsofia Zvolenszky (eds.), *Semantics and Linguistic Theory (SALT) 11*, 231–247. Cornell University, Ithaca, NY: CLC Publications.

Gualmini, Andrea, Sarah Hulsey, Valentine Hacquard & Danny Fox. 2008. The Question-Answer Requirement for scope assignment. *Natural Language Semantics* 16. 205–237.

Gualmini, Andrea & Bernhard Schwarz. 2009. Solving the learnability problems in the acquisition of semantics. *Journal of Semantics* 26. 185–215.

Huang, Yi Ting, Elizabeth Spelke & Jesse Snedeker. 2013. What exactly do numbers mean? *Language Learning and Development* 9. 105–129.

Ivlieva, Natalia. 2013. *Scalar implicatures and the grammar of plurality and disjunction*: Massachusetts Institute of Technology dissertation.

Katsos, Napoleon & Dorothy V.M. Bishop. 2011. Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition* 120(1). 67–81.

Krifka, Manfred. 1995. Common nouns: A contrastive analysis of English and Chinese. In Greg N. Carlson & Francis Jeffrey Pelletier (eds.), *The generic book*, 398–411. Chicago: Chicago University Press.

Lazaridou-Chatzigoga, Dimitra, Napoleon Katsos & Linnaea Stockall. 2015. Genericity is easy? Formal and experimental perspectives. *Ratio: An international journal of analytic philosophy* 28(4). 470—494.

Lewis, Shevaun. 2013. *Pragmatic enrichment in language processing and development*: University of Maryland dissertation.

Magri, Giorgio. 2014. An account for the homogeneity effects triggered by plural definites and conjunction based on double strengthening. In Salvatore Pistoia Reda (ed.), *Pragmatics, semantics and the case of scalar implicatures*, 99–145. Palgrave Macmillan.

Marchman, V.A., K. Plunkett & J. Goodman. 1997. Overregularization in English plural and past tense inflectional morphology: A response to Marcus (1995). *Journal of Child Language* 24(3). 767–779.

Marcus, G.F., S. Pinker, M. Ullman, M. Hollander, T.J. Rosen & F. Xu. 1992. *Overregularization in language acquisition*, vol. 57 (Monographs of the Society for Research in Child Development 4, Serial No. 228).

Mayr, Clemens. 2015. Plural definite dps presuppose multiplicity via embedded exhaustification. In *Proceedings of SALT 25*, .

Mervis, Carolyn B. & Kathy E. Johnson. 1991. Acquisition of the plural morpheme: A case study. *Developmental Psychology* 27. 222–235.

Miller, K., C. Schmitt, H. Chang & A. Munn. 2005. Young children understand some implicatures. In *Proceedings of the 29th annual boston university conference on language development*, Cascadilla Press.

Noveck, Ira. 2001. When children are more logical than adults: experimental investigations of scalar implicatures. *Cognition* 78(8). 165–188.

Papafragou, Anna & Julien Musolino. 2003. Scalar implicatures: experiments at the semantics–pragmatics interface. *Cognition* 86(3). 253–282.

Patson, N.D., G.E. George & T. Warren. 2014. The conceptual representation of number. *Quarterly Journal of Experimental Psychology* 67. 1349–1365.

Patson, Nikole D. 2016. Evidence in support of a scalar implicature account of plurality. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 42(7). 1140–1153.

Pearson, Hazel, Manizeh Khan & Jesse Snedeker. 2011. Even more evidence for the emptiness of plurality: An experimental investigation of plural interpretation as a species of scalar implicature. In David Lutz & Nan Li (eds.), *Semantics and Linguistic Theory (SALT) 20*, Cornell University Ithaca, NY: CLC Publications.

Roberts, Craige. 1996. Information structure: Towards an integrated formal theory of pragmatics. In J. H. Yoon & A. Kathol (eds.), *Osuwpl volume 49: Papers in semantics*, 1–54. The Ohio State University Department of Linguistics.

Sauerland, Uli. 2003. A new semantics for number. In Robert B. Young & Yuping Zhou (eds.), *Semantics and Linguistic Theory (SALT) 13*, 258–275. Cornell University Ithaca, NY: CLC Publications.

Sauerland, Uli. 2005. DP is not a scope island. *Linguistic Inquiry* 36(2). 303–314.

Sauerland, Uli, Jan Andersen & Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In Stephan Kepser & Marga Reis (eds.), *Linguistic evidence: Empirical, theoretical, and computational perspectives*, 413–434. Mouton de Gruyter.

Singh, Raj, Ken Wexler, Andrea Astle-Rahim, Deepthi Kamawar & Danny Fox. 2016. Children interpret disjunction as conjunction: Consequences for theories of implicature and child development. *Natural Language Semantics* 24(4). 305–352.

Sneed, Elisa. 2005. The role of input in the acquisition of generic NPs. In Alejna Brugos, Manuella R. Clark-Cotton & Seungwan Ha (eds.), *BUCLD 29 online proceedings supplement*, .

Sneed, Elisa. 2008. *Input in the acquisition of genericity*. Evanston, Illinois: Northwestern University dissertation.

Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: On higher-order implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and implicature in compositional semantics*, 243–281. Palgrave.

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

van Tiel, Bob, Emiel van Miltenburg, Natalia Zevakhina & Bart Geurts. 2014. Scalar diversity. *Journal of Semantics* doi:10.1093/jos/ffu017.

Tieu, Lyn, Cory Bill, Jacopo Romoli & Stephen Crain. 2014. Plurality inferences are scalar implicatures: Evidence from acquisition. In Todd Snider (ed.), *Proceedings of the 24th semantics and linguistic theory conference*, 122–136.

Tieu, Lyn, Cory Bill, Jacopo Romoli & Stephen Crain. 2015. Plurality inferences are scalar implicatures: Further evidence from acquisition. Paper presented at the Workshop on Experimental and Crosslinguistic Evidence for the Distinction between Implicatures and Presuppositions.

Tieu, Lyn, Cory Bill, Jacopo Romoli & Stephen Crain. Under review. Testing theories of plural meanings: Insights from child language.

Tieu, Lyn, Jacopo Romoli, Peng Zhou & Stephen Crain. 2016. Children's knowledge of free choice inferences and scalar implicatures. *Journal of Semantics* 33(2). 269–298.

Winitz, H., R. Sanders & J. Kort. 1981. Comprehension and production of the /schwa z/ plural allomorph. *Journal of Psycholinguistic Research* 10(3). 259–271.

Winter, Yoad. 2001. *Flexibility principles in boolean semantics*. Cambridge, Mass.: MIT Press.

Wood, Justin N., Sid Kouider & Susan Carey. 2009. Acquisition of singular-plural morphology. *Developmental Psychology* 45(1). 202–206.

Zapf, Jennifer A. & Linda B. Smith. 2008. Meaning matters in children's plural productions. *Cognition* 108(2). 466–476.

Zondervan, Arjen. 2007. Effects of Question Under Discussion and focus on scalar implicatures. In M.E. Kluck & E.J. Smits (eds.), *Proceedings of the fifth semantics in the netherlands day*, 39–52.

Zondervan, Arjen. 2009. Experiments on qud and focus as a contextual constraint on scalar implicature calculation. In Uli Sauerland & Kazuko Yatsushiro (eds.), *From experiment to theory, presuppositions, negation and scalar implicatures*, Palgrave.

Zondervan, Arjen. 2010. *Scalar implicatures or focus*: Utrecht Institute of Linguistics dissertation.

Zweig, Eytan. 2009. Number neutral bare plurals and the multiplicity agreement. *Linguistics and Philosophy* 32. 353–407.