On the nature of the plurality inference: Ladybugs for Anne[*]

Penka Stateva, Sara Andreetta and Arthur Stepanov
Center for Cognitive Science of Language, University of Nova Gorica

1. Introduction

Recent explorations in formal pragmatics provide ample argumentation and some experimental support to the claim that plural nominal morphology does not lexically encode plurality but rather triggers a plurality inference of a pragmatic nature (Sauerland 2003, Sauerland, Andersen & Yatsushiro 2005; Spector 2007; Sauerland 2008; Magri 2014, Ivlieva 2013, Tieu et al 2014, Mayr 2015). This claim is based on the observation that plural nominal interpretations do not always entail a reference to more than one individual. Rather, there are certain well-defined configurations in which the meaning of plural nominal expressions is either consistent with an 'at least one' reading or even requires that reading. Among these, most discussed are bare plurals in the scope of negation, as in (1b), which disallow the 'more than one' interpretation of the same plural in the non-negated counterpart of (1b) in (1a).

(1) a. For his experiment, Peter recruited **subjects** from the chess club.

b. For his experiment, Peter didn't recruit **subjects** from the chess club.

[*] To an amazing scholar who has given a superb example of academic excellence, professionalism, vision and all the valuable qualities that we can strive to follow. Happy Birthday, Anne! This research was supported in part by the large-scale project AThEME (Grant Agreement 613465) from the European Union's Seventh Framework Programme, and by the Slovenian Research Agency program No. P6-0382. Any errors and inconsistencies are the authors' responsibility.

The contrast in (1) highlights what is explicitly argued for in Sauerland (2003) that the plural nominal morpheme is semantically unmarked, i.e. plural noun interpretations are consistent with singular or plural references. Other, even more intriguing examples involve the interpretation of bare plurals or narrow-scope existentially quantified plural nominals in the scope of a universal quantifier, as in (2). Both of these examples have been claimed not to require a 'more than one' reading for the respective plural nominals or rather, to allow for exceptions with respect to plurality.

(2) a. Every student should invite his **sisters**. (Sauerland 2008)

b. Yesterday, every student solved (some) **difficult problems**. (Spector 2007)

More specifically, Sauerland (2008) claims that (2a) is judged true not only in contexts in which every student has several sisters but also if some (but not all) students do not have a plurality of sisters. Spector (2007) describes similar sentences as being genuinely ambiguous. In his view, (2b) has an interpretation according to which every student solved several difficult problems and another weaker reading (coinciding with the one described in Sauerland 2008) according to which some students solved several difficult problems while the rest could have solved only one.

Both Sauerland (2008) and Spector (2007) suggest theories of nominal plural morphology according to which reference to pluralities in interpretations of plural nominals result from pragmatic reasoning. However, the suggested triggering mechanisms are based on the involvement of different pragmatic strategies. How exactly sentences like (2) are evaluated is an empirical question whose answer can help tease the two theories apart. Our goal is to offer some experimentally supported conclusions that could enhance such a discussion and contribute to the research about the nature of the plurality inference. To this end, in this paper we report results from two experiments in Italian that investigate the processing of plural nominal expressions in the scope of quantifiers. We start with a brief presentation of the two

competing theories and their predictions regarding interpretations in embedded contexts in Section 2. Section 3 introduces the design, materials, participants and results of Experiments 1 and 2. We conclude in Section 4 with a general discussion of the results. We consider the relevance of the number paradigm to the two competing pragmatic principles, and consequently to the notions of implicated presupposition and second-order implicature which will be introduced shortly.

2. Ways of deriving the plurality inference

2.1 Non-singularity as an instance of implicated presuppositions

Chierchia (1998) cites (3) which is due to Krifka (p.c.) who suggests that plural morphology on nouns that are modified by numerals is only a result of an agreement process.

(3) The average Italian family buys 0.5 cars (/*car) per year.

Sauerland (2003, 2008) and Sauerland, Andersen & Yatsushiro (2005) push the idea further by arguing that generally the plural nominal morpheme is not lexically specified to contribute to the 'more than one' interpretation of plural nominals. The (informal) meaning of a sentence like (4a) is roughly a result of coupling (4b) which is the compositionally derived interpretation (literal meaning) of (4a) including the presupposition of the possessive and what is termed as *implicated presupposition* by Sauerland or *antipresupposition* by Percus (2006) in (4c):

(4) a. Mary's sisters are bright.

b. Mary has one or more sisters and they are bright.

c. It is not the case that Mary has exactly/only one sister.

In effect, (4c) which is responsible for the plural interpretation of *sisters* in (4a) results from an alternative evaluation procedure triggered by the application of the pragmatic principle *Maximize Presupposition* originally due to Heim (1991). Speakers

are claimed to observe *Maximize Presupposition* by evaluating truth-conditionally equivalent alternatives in a given context whose presuppositions are satisfied in that context and choosing the one with maximal presuppositional content. Alternatives sharing the same assertive content would differ with respect to presupposed content if one of them contains a conventional presupposition trigger that the other one lacks. Crucially, since weaker presuppositional alternatives are relevant for evaluation alongside stronger alternatives, uttering a weaker alternative under the assumption that *Maximize Presupposition* is observed triggers an inference that the stronger alternative's presupposition is not satisfied. A presuppositionally stronger alternative to (4a) is (5) in which the noun sister is marked with the semantic feature *singular*.

(5) Mary's sister is bright.

That feature is assumed to presuppose singularity of the nominal phrase referent. Preference of (4a) over (5) triggers (4c) and respectively the speaker's meaning of (4a) that Mary has more than one sister and her sisters are bright.

The suggested mechanism of deriving implicated presuppositions is strongly reminiscent of the mechanism of deriving implicatures, and more specifically, of scalar implicatures which can save an utterance from an apparent violation of the *Maxim of Quantity* when weaker alternatives, i.e. asymmetrically entailed alternatives, are preferred over stronger ones. This implies a possibility that *Maximize Presupposition* be eliminated as an operative pragmatic principle and implicated presuppositions be analyzed as a kind of implicatures modulo some additional assumptions that widen the domain of the *Maxim of Quantity*.[1,2] However, Sauerland (2008) argues that the mechanism of deriving implicated presuppositions cannot be reduced to that of deriving scalar implicatures because the two kinds of inferences

[1]This possibility has been explored in detail in Singh (2011), Schlenker (2012), Magri (2013).

[2]Our goal here is not to take a stand on the question whether implicated presuppositions are reducible to scalar implicatures. Rather, we would like to test some predictions of two competing analysis of deriving the plurality inference related to plural nominal morphology.

differ with respect to a number of properties, and in fact, implicated presuppositions share some features with presuppositions, and some with implicatures which makes them a unique class.[3] It follows then, that if the <singular, plural> opposition is part of the set that is relevant to *Maximize Presupposition* it will manifest the properties ascribed to the whole class of oppositions that are linked to that principle and, respectively, to implicated presuppositions. This class standardly includes oppositions like <know, believe>, <the, a>, <both, every/all>, <the, every/all> among few other.

One of the properties of implicated presuppositions that Sauerland (2008) identifies regards projection through a universal quantifier. Sauerland's analysis predicts that an implicated presupposition trigger in the scope of a universal quantifier would not distribute that inference universally. Within his framework, deriving an implicated presupposition is a post-grammatical phenomenon. It follows that if *p* is the conventional presupposition of *q*'s alternative *s* with the same assertive content but a stronger presuppositional load, then 'not *p*' is *q*'s implicated presupposition. Consequently, if the presupposition trigger is conventional, then *s* will presuppose that every individual in the restriction of the quantifier will have a property that satisfies the presupposition *p*. On the other hand, *q*'s implicated presupposition 'not *p*' would entail that not every individual in the quantifier's restriction will have the property that satisfies the presupposition of *s*. But since $\neg\forall xP(x)$ is logically equivalent to $\exists x\neg P(x)$ it follows that anti-presuppositions that spur from triggers in the scope of a universal quantifier do not project freely through it, though they are not cancelled and in that sense, universal quantifiers are not barriers for anti-presupposition projection. Let us illustrate the point with some

[3]For example, implicated presuppositions are claimed to have a weaker epistemic status from both presuppositions and implicatures. With respect to projection through negation, they are claimed to pattern with presuppositions and differ from implicatures.

examples. Consider (6a) and (6b) which qualify as alternatives in the sense of *Maximize Presupposition*:

(6) a. Every guest knows that his presence was appreciated.

b. Every guest believes that his presence was appreciated.

In the context of evaluation (6a) presupposes that every guest's presence was appreciated and it is, therefore, inappropriate to continue it with (7) under the assumption that Mark was one of the guests:

(7) But Mark's presence was not appreciated.

Since (6b) is a weaker presuppositional alternative of (6a), its utterance is paired with the anti-presupposition that it is not true that every guest's presence was appreciated. (6b) is then felicitous in contexts in which at least some (and possibly all) guest's presence was not appreciated so (7) can be an appropriate continuation of (6b). Importantly, from (7) one can infer that there were guests whose presence was appreciated so the theory seems to predict correctly that every is indeed not a barrier for anti-presupposition projection but allows for exceptions in distributing the relevant property to members of its restriction.

The discussion above makes the difference between implicated presuppositions and conventional presuppositions obvious: while the latter project through the universal quantifier, the former do not project freely in that way. Comparing implicated presuppositions and implicatures, on the other hand, with respect to this property is not trivial. The conclusion depends on whether the compared theory of scalar implicatures assumes the possibility that they can also be derived locally.[4] Very

[4]Scalar implicatures are analyzed as global inferences by the Gricean and earlier neo-Gricean pragmatic frameworks. On the other hand, the grammatical theory of implicatures (Chierchia 2004, Fox 2007, Chierchia, Fox and Spector 2012, Chierchia 2013) which assumes that they are derived as part of the truth-conditional content by freely adding an operator that negates contextually added alternatives allows for their local, as well as, global derivation.

informally, (8a) is associated with the globally derived implicature in (8b) and the locally derived one (which entails (8b)) in (8c):

(8) a. Every child ate some of the chocolate desert on her plate.

b. Not every child ate all of the chocolate desert on her plate.

c. Every child ate some but not all of the chocolate desert on her plate.

It is obvious that local implicatures, similarly to conventional presuppositions freely project through the universal quantifier. Therefore, one cannot continue (8c) with (9):

(9) But even crumbs were not to be found on Mary's plate.

Contrary to that, (9) is an appropriate continuation of (8b). This implies that global implicatures are similar to implicated presuppositions with respect to projection across the universal quantifier. In a similar vein, these inferences are not cancelled by the universal quantifier but the force of the implicature-based quantifier is existential. To illustrate this with (8b): (8b) describes contexts in which at least one child did not eat all of the desert on her plate. To summarize the discussion, implicated presuppositions, which were shown to differ from conventional presuppositions, differ also from implicatures with respect to their projection properties because they only pattern with globally derived implicatures but not with locally derived ones. However, it must be mentioned that if *Maximize Presupposition* is redefined so that it can also apply locally[5], then scalar implicatures and implicated presuppositions would not exhibit any difference with respect to projection: we would then be able to say that globally derived implicated presuppositions pattern with globally derived implicatures, and locally derived implicated presuppositions pattern with locally derived implicatures.

[5]See references in ft.1 and also Percus (2006).

Returning to the inference related to plural nominal morphology which is the focus of this paper, let us consider Sauerland's theory's predictions about the anti-singularity inference accompanying sentences with plural nominals. In which kinds of contexts can one use (10) felicitously?

(10) Every ladybug has (some) dots.

If the anti-singularity inference that is triggered by the plural dots is to be analyzed as an instance of implicated presuppositions, then the speaker's interpretation of (10) results from putting together (11a) and (11b).

(11) a. Every ladybug has one or more dots.

b. It is not the case that every ladybug has a single dot

(at least one ladybug has more than one dot).[6]

It follows that under the assumptions of the theory of implicated presuppositions, (10) should be equally acceptable in contexts in which some ladybugs from a relevant group have only one dot irrespective of the number of exceptions. What is required in this case is that there is at least one ladybug with several dots. Also, (11b) is consistent with a context in which all ladybugs have several dots.

The predictions change, if *Maximize Presupposition* could be applied locally. While the global implicated presupposition would have the form in (12a), the local one would look like (12b):

(12) a. $\neg\{x: x \text{ is a ladybug}\} \subseteq \{x: x \text{ has exactly one dot}\}$

b. $\{x: x \text{ is a ladybug}\} \subseteq \{x: x \text{ does not have exactly one dot}\}$

When (12) is coupled with (11a), we predict that if *Maximize Presupposition* is applied only locally, then (10) will be acceptable just in those contexts in which every

[6]We side with Spector (2007) that more details have to be made explicit in the mechanism of deriving implicated presuppositions. He warns that it might be problematic to simply negate the stronger presupposition since the resulting inference might turn out to be weaker than necessary.

ladybug has several dots. If *Maximize Presupposition* can apply locally, as well as globally, then (10) is not an appropriate case to distinguish that theoretical option from the option of applying *Maximize Presupposition* only globally, as in the original theory of Sauerland, since the predictions of both options coincide given that (12b) entails (12a).[7]

Out of the set of potential quantificational 'barriers' for projection of implicated presuppositions Sauerland (2008) has singled out only the universal quantifier. However, extension to other quantifiers like the non-monotone '*exactly n NP*' suggests that they function in a similar manner.[8] Consider (13a) and its truth-conditional meaning in (13b) without the implicated presupposition.

(13) a. Exactly three ladybugs have dots.

b. Exactly three ladybugs have one or several dots.

A globally derived non-singularity implicated presupposition as in (14a) complements (13b) to make (13a) acceptable in contexts in which zero, one or two ladybugs out of a group that contains exactly three dotted ladybugs have exactly one dot.

(14) a. $\neg |\{x: x \text{ is a ladybug}\} \cap \{x: x \text{ has exactly one dot}\}| = 3$

b. $|\{x: x \text{ is a ladybug}\} \cap \{x: x \text{ does not have exactly one dot}\}| = 3$

On the other hand, (14b) which instantiates a local implicated anti-singularity presupposition leads to interpreting (13a) as requiring that exactly three ladybugs in a relevant group of ladybugs are dotted and they all have several dots.

[7]The line of reasoning and consequently design of the experiment is influenced by reasoning about scalar implicature derivation in Chemla and Spector (2011).

[8]Sauerland et al (2005) have a special proviso for applying *Maximize Presupposition* to the scope of existential quantifiers (a principle stipulating potential exceptions for the application of *Maximize Presupposition*). We believe that *Maximize Presupposition* is not prevented from applying in (13a).

2.2 Non-singularity as an instance of higher-order implicatures

A defining feature of Sauerland's theory of nominal plural depends on the assumption that the plural morphology is semantically non-marked, while singular morphology contributes only presuppositionally by allowing just atomic individuals in the set denoted by the noun to which it is attached. It is therefore, possible to see sentences that only differ with respect to nominal morphology as relevant alternatives in the sense of *Maximize Presupposition* and derive the 'at least two'-reading of plural nominals though reasoning based on its application. A different pragmatic-based path is taken by Spector (2007) who suggests that the plurality inference is a special kind of scalar implicature.[9] But in order to derive that inference of (15a), for example, Spector's theory has to (i) assume contra Sauerland but in line with Krifka (1989) and others that plural nominal morphology is semantically specified as a function that applies to a set of atomic individuals to give a join-semilattice and (ii) present (15b) as its legitimate, more informative scalar alternative which is apparently not trivial since scalar alternatives are supposed to have the same brevity, and level of lexicalization (cf. Atlas and Levinson 1981).

(15) a. The ladybug has dots.

b. The ladybug has exactly one dot.

Spector reasons that if there are arguments to support the status of (15b) as a scalar alternative to (15a), then negating it derives the 'at least two' inference with which the basic semantic reading of (15a) is enriched.

To present the gist of the theory in some details, let us start with the assumption in (i). Under it, the plural noun *dots* denotes a set of both atomic and non-atomic individuals so the compositional meaning of (15a) is as in (16):

[9]That proposal is orthogonal to the discussion about the potential elimination of *Maximize Presupposition* as an operative pragmatic principle.

(16) $\exists X[\forall x[(\text{atomic}(x) \wedge x << X) \rightarrow x \in [[\text{dot}]]] \wedge$ The ladybug has X]

(16) describes contexts in which the relevant ladybug has one or more dots. That semantic treatment of the plural marker makes obvious the logical relation between (15a) and its counterpart with a singular nominal predicate in (17a) and its literal meaning in (17b).

(17) a. The ladybug has a dot.

b. $\exists x(x \in [[\text{dot}]] \wedge$ The ladybug has x)

On the one hand (17a) entails (15a) which makes the pair of alternatives based on the opposition <pl,sg> legitimate. On the other hand, when (16) is satisfied, so is (17b) and as noted by Spector, we cannot claim that (17a) entails (15b) asymmetrically. However, these two sentences do not really have the same meaning in terms of pragmatic strengthening. For one, (17) implies (15b) and the reason for that is that there is one more pragmatic scale that is activated in the evaluation of (17a) and it is based on the opposition <a, several>. Since (18) asymmetrically entails (17a), its negation is implicated upon uttering (17a).

(18) The ladybug has several dots.

As a result, we get (15b) as the pragmatically strengthened meaning for (17a). The cornerstone of Spector's proposal regards in general the possibility to widen the set of scalar alternatives that are available for exhaustification so that it also includes not only propositions that correspond to literal interpretations but also to their pragmatically enriched counterparts. Since (15b)'s content is identical to the pragmatically strengthened proposition expressed by (17a), and (17a) is a scalar alternative to (15a), then (15b) is also under consideration along with (15a). In this way, the intuitively correct path to consider (15a) and (15b) as scalar alternatives is theoretically substantiated. The speaker's meaning of (15a), in accord with intuition,

is derived as a conjunction of (16) and the negation of the proposition expressed by (15b).

Both theories we considered derive correctly the non-singularity inference related to the interpretation of plural nominals. Let us now consider whether their predictions differ with respect to the interpretation of plural nominals in the scope of the universal quantifier, and the non-monotone quantifier *exactly three*, respectively. Consider (10) again, repeated here as (19a) and its alternative in (19b) resulting from the opposition <pl, sg>:

(19) a. Every ladybug has (some) dots.

b. Every ladybug has a dot.

Given Spector's assumptions about the literal meaning of the plural, (19a) and (19b) are synonymous with respect to their literal meaning in (20).

(20) Every ladybug has one or more than one dots.

Given the opposition <a, several>, (19b)'s literal meaning can be strengthened by adding to (20) the negation of (19b)'s alternative in (21).

(21) Every ladybug has several dots.

The pragmatically enriched meaning of (19b) is then equivalent to (22).

(22) = (20) & ¬ (21)

= Every ladybug has one or more dots and not every ladybug has several dots.

Since Spector's proposal allows for iterative exhaustification of alternatives in computing the speaker's meaning of (19a), now we can evaluate (22) as an alternative of its literal meaning in (20) since this is the already strengthened proposition expressed by (19b). As a result we predict (19a) to express the proposition in (23):

(23) = (20) & ¬ (22)

= Every ladybug has one or more dots and it is not the case that every ladybug has one or more dots and not every ladybug has several dots

= Every ladybug has one or more dots and it is not the case that not every ladybug has several dots

= Every ladybug has one or more dots and every ladybug has several dots

= Every ladybug has several dots.

This predicted meaning for (19a) differs from the meaning predicted for it by Sauerland's theory. As we saw above, the latter predicts possible ladybug exceptions from the 'at least two' requirement about the number of their dots. However, this is not an unwelcome result, according to Spector who argues that (19a) is genuinely ambiguous between what he calls a strong meaning (cf. (23)) and a weak one which coincides with Sauerland's only predicted interpretation for that sentence.[10] The source of ambiguity is not structural but pragmatic in nature. As Spector correctly points out, (19a) contains more than one scalar item, and potentially more alternatives should be considered if the quantificational determiner *every* is viewed as a source of alternative evaluation.[11] If we consider all scalar alternatives based on the oppositions <pl, sg>, <a, several>, and <every, some>, all of the sentences in (24) are available for evaluating the strengthened meaning of (19a):

(24) a. Every ladybug has dots.

b. Every ladybug has a dot.

c. Every ladybug has several dots.

[10]Spector relays a personal communication with U. Sauerland which focuses on the fact that assuming a genuine ambiguity in (19a) might be considered problematic given that one of the suggested readings entails the other. We side with Sauerland that this is a potential concern.

[11]We voice a concern already mentioned by Spector that a theory of pragmatic ambiguity needs to be restricted on principled grounds and this deserves further research.

d. Some ladybug has dots.

e. Some ladybug has a dot.

f. Some ladybug has several dots.

Since the pairs (24a-b) and (24d-e) have the same literal meaning and the second pair is entailed by the first pair of sentences, after the first round of strengthening (exhaustification) (24b) is pragmatically enriched as in (25):

(25) = (24b) & ¬ (24c) &¬ (24f)

= Every ladybug has one or more than one dot and not every ladybug has more than one dot and it is not the case that some ladybugs have more than one dot

= Every ladybug has exactly one dot.

This is the only pragmatically enriched alternative to (19a) whose negation is consistent with the literal meaning of (19a) and not entailed by it. So strengthening that literal meaning amounts to (26):

(26) Every ladybug has one or more dots and it is not the case that every ladybug has exactly one dot.

This is what Spector calls the weak reading of (19a) and the only reading predicted by Sauerland´s theory. Finding evidence for or against the postulated ambiguity is obviously extremely difficult since contexts that verify (19a) on the strong reading are also contexts that verify it on the weak reading. However, the predictions of the two theories go astray with respect to the interpretations of the plural nominals in the scope of non-monotone quantifiers. (13a), repeated as (27) is a case at hand.

(27) Exactly three ladybugs have dots.

Since the non-monotone quantifier in (27) is not a member of scalar oppositions, then the only alternatives that are under consideration for strengthening the literal meaning of (27) are due to the plural morpheme in *dots*. Its direct alternative in (28a) can be strengthened by negating (28b) and deriving (28c):

(28) a. Exactly three ladybugs have a dot.

b. Exactly three ladybugs have several dots.

c. Exactly three ladybugs have one or more than one dot and it is not the case that exactly three ladybugs have more than one dot

= Exactly three ladybugs have one or more than one dot and at least one ladybug has exactly one dot.

To derive the strengthened meaning of (27) we combine its literal meaning and the negation of the proposition that expresses the strengthened meaning of (28a). Thus we predict, within the framework of higher-order implicatures that the speaker's meaning of (27) is as in (29):

(29) Exactly three ladybugs have one or more than one dot and it is not the case that exactly three ladybugs have one or more than one dot and at least one ladybug has exactly one dot

= Exactly three ladybugs are dotted and exactly three ladybugs have more than one dot.

We thus see that the theory of plural nominal morphology based on the application of *Maximize Presupposition* and the theory of higher-order implicatures make different predictions with respect to interpreting plural nominals in the scope of the universal quantifier and the scope of non-monotone quantifiers. Our goal is to provide some experimental basis for evaluating these predictions.

1 3. How many dots satisfy the predicate have/has dots?

3.1. Laying out predictions for an experimental study

The goal of Experiments 1 and 2 reported below is to test the predictions made by the theory based on *Maximize Presupposition* as outlined in Sauerland (2008) and Sauerland et al (2005) the theory of higher-order implicatures, with respect to the plural morphology as outlined in the previous section. Under both the *Maximize Presupposition* theory and the theory of higher-order implicatures, we expect speakers to evaluate sentences with unembedded plural nominals like *Some dots are red* as true and felicitous if the context contains a plurality of dots. Consequently, we expect significantly lower acceptance if only one dot in the context is red.

These theories, however, make different predictions with respect to plurals in embedded contexts. Specifically, we consider two cases of such embedding. The first case concerns plurals in the scope of the universal quantifier, in a sentence like (10) repeated here:

(10) Every ladybug has some dots.

With respect to (10), the *Maximize Presuppositions* theory predicts that there be no significant difference in acceptance between contexts in which each ladybug has a plurality of dots, and contexts in which some ladybugs have only one dot, as long as there exists at least one ladybug with more than one dot. A significant difference in acceptance is expected between contexts in which all ladybugs have only one dot, on the one hand, and contexts containing at least one multiply-dotted ladybug, on the other. The same outcome is expected under the higher-order implicatures theory. These predictions are reflected in Figure 1:

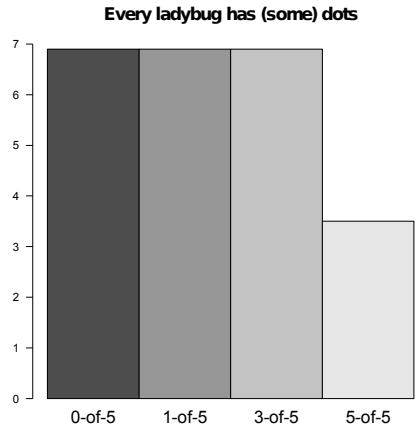


Figure 1. Sauerland's (2008) prediction and Spector's (2007) prediction regarding the weak reading of sentence (10)

At the same time, since the latter theory entails a genuine ambiguity for (10), it predicts that individual subjects might have preferences for the weak or the strong reading. If they accept predominantly the strong reading then the higher-order implicature theory predicts a significant difference in acceptance between contexts in which all ladybugs have more than one dot and contexts where some, and possibly all, ladybugs have only one dot. This is illustrated on Figure 2:

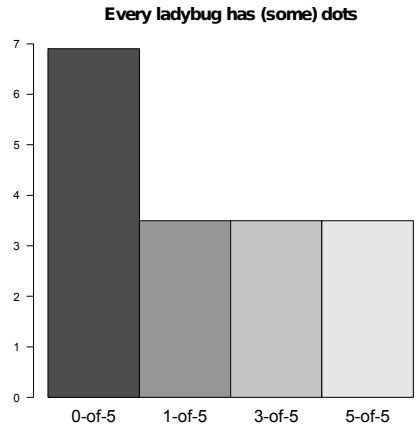


Figure 2. Spector's (2007) prediction regarding the strong reading of sentence (10)

The two theories also make different predictions with respect to plurals embedded under non-monotone quantifiers like in (13a) repeated here:

(13) a. Exactly three ladybugs have some dots

The *Maximize Presupposition* theory predicts (13a) to be acceptable in contexts in which exactly three ladybugs in the relevant group of, say, five are dotted and at least one of them has more than one dot. We therefore expect no significant difference in acceptance between contexts which satisfy the condition of having exactly three dotted ladybugs, but, within that subset, there are 0, 1 or 2 ladybugs that only have 1 dot. At the same time, we expect all of these contexts to be evaluated significantly better than contexts in which three ladybugs have only one dot. See Figure 3:

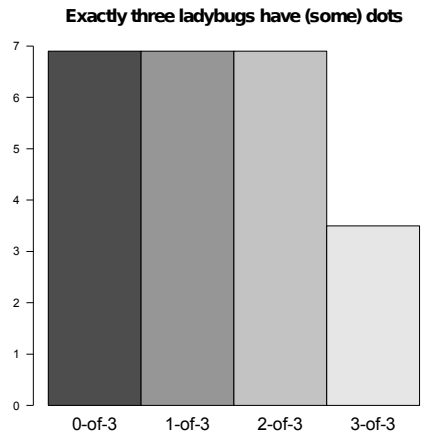


Figure 3. Sauerland's (2008) prediction regarding sentence (13a)

The higher-order implicature theory, however, predicts that in this case only contexts with three multiply-dotted ladybugs are acceptable. A significant difference in acceptability is expected between contexts with 0 one-dotted ladybugs and contexts with 1, 2, or 3 ladybugs with one dot in the relevant set. See Figure 4:

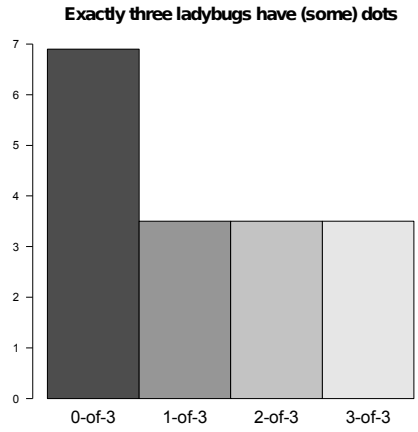


Figure 4. Spector's (2007) prediction regarding sentence (13a)

3.2. Experiment 1

3.2.1. Design and materials

Experiment 1 had two parts. One part involved evaluation of sequences of red, blue and green dots (the "red dots" part), and another part involved evaluating pictures of dotted ladybugs (the "ladybugs" part). The stimuli for both parts contained a visual context and a sentence that accompanied that context.

The factor manipulated in the "red dots" part of the experiment was the number of red dots (specifically, one or two), as a proportion of the total number in a sequence of dots. An additional manipulated factor was the color of the non-red dots (green, blue). The "dots" set contained 26 conditions of which 10 were target conditions. All target conditions were accompanied by the Italian sentence in (30):

(30) Alcuni cerchi sono rossi

`Some dots are red.`

The set of contexts for the target sentence included 2 contexts with 3 dots, 4 contexts with 5 dots, and 4 contexts with 7 dots. Half of these contained 1 red dot and the other half contained 2 red dots. Each line of dots was completed either by green or by blue dots and the configurations of adding blue or green were matched. Only one item was created per each "number of dots per color" configuration. Some examples are given below:

(31) a. 

1/3

b. 

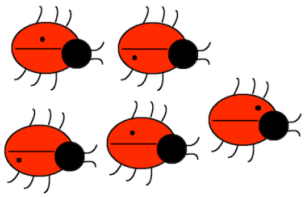
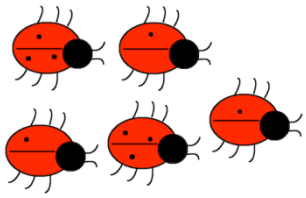
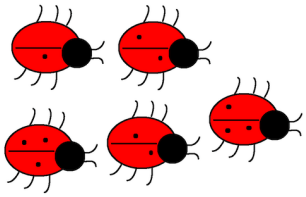
2/5

c. 

1/7

The resulting 10 context-sentence items were supplemented with 16 contexts that served as controls and/or fillers and used dot sequences constructed of the same three colors in different combinations. Accompanying sentences used the same sentence frame but a different determiner: *many*, *few*, *most*, *all*, *three/four*, *no*, *half*. Sentences containing *many* and *few* were not used as controls, but only as fillers.

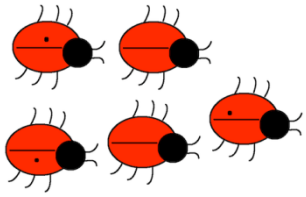
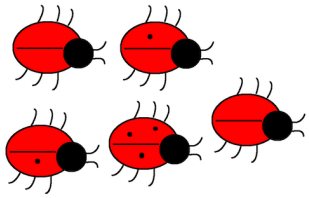
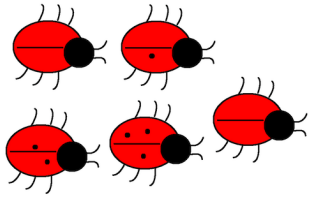
The factor manipulated in the "ladybugs" part of the experiment was the number of single-dotted ladybugs in a sequence of 5 (five) ladybugs. The "ladybug" set of materials contained 18 conditions of which 8 were target conditions. Specifically, in Condition 1, each ladybug had exactly 1 dot, as shown in (32a). In Condition 2, 3 ladybugs out of 5 had one dot, while the rest have at least 2 dots (32b). In Condition 3, only one ladybug had 1 dot (32c). Condition 4 had no single-dotted ladybugs. Conditions 1-4 were accompanied by the sentence in (32d).

(32) a.  b.  c. 

d. Ogni coccinella ha alcuni puntini.

`Every ladybug has some dots.`

The other four conditions were slightly different. In each condition, a picture of 5 ladybugs was presented, of which 3 were dotted and the other 2 not. Specifically, in condition 5, each of the 3 dotted ladybugs had 1 dot. In condition 6, 2 out of the 3 dotted ladybugs had 1 dot while the third had 3 dots. In condition 7, only 1 out of the 3 dotted ladybugs had 1 dot while the rest had at least 2. In condition 8, there were no single-dotted ladybugs. These conditions were accompanied by the sentence in (33d):

(33)a.  b.  c. 

d. Esattamente tre coccinelle hanno alcuni puntini.

´Exactly three ladybugs have some dots.´

Two items were created per each of the above eight conditions, making it the total of 16 target items.

Sentences (32d) and (33d) were also used in contexts in which they were either unequivocally true and felicitous (for example, each ladybug had at least two dots, or exactly three ladybugs had at least two dots and the rest were not dotted) or definitely false (for example, 4 ladybugs had at least two dots but 1 was not dotted, or only two out of the 5 ladybugs had any dots while the others were not dotted). The number of true and false answers per quantifier per color was matched in the set.

The additional 12 items served as controls. They all contained pictures of sets of 5 ladybugs and were accompanied by sentences similar to the ones used in the targets but nominals that they had were embedded under quantifiers such as *some*, *no*, *every* and *exactly three*. The plural nominal expressions were all of the form *at least/exactly/at most n dots*. An example is given in (34):

(34) Nessuna coccinella ha più di quattro puntini.

´No ladybug has more than 4 dots.´

The two sets of stimuli were presented in different order in two versions of the experiment. One version presented the "dots" set first while the other presented the "ladybugs" set first. The distribution of the two versions was balanced across participants.

3.2.2. Participants

34 self-reported monolingual Italian adults participated in this study (14 female). The participants were aged between 24-39 years old (mean age 31,1 years, standard deviation = 3,2). They were recruited through personal contacts in the area of Udine, Italy. A short online questionnaire preceding the experiment contained information about age, native language, vision and potential color-blindedness.

3.2.3. Procedure

The participants were presented with a series of visual contexts along the lines described above, and an accompanying sentence, on the same screen. The participants were asked to evaluate how well the sentence fits the visual context and deliver a judgment using a Likert scale from 1 (definitely inappropriate) to 7 (definitely appropriate). Stimuli were presented using the Ibex farm platform (by Alex Drummond, http://spellout.net/ibexfarm/). Each subject received all conditions but presentation was fully randomized per participant within each set.

3.2.4. Results

We excluded the data from 6 participants who correctly evaluated less than 76% of the control sentences. In particular, we considered as mistakes the following two types of answers: i) definitely failed control, in cases when the participant answered 1 or 2 when the correct answer was 6 or 7 (and reciprocally 6 or 7 in the cases of false controls), and ii) the deviation from standard, that is, if a participant answered 3, 4 or 5 when the correct answer was 1 or 2, or 6 or 7, respectively. This left the cleaned data from 28 participants to be used in further analyses. Of these 28 participants, 15 saw the "ladybugs" set first, and 13 who saw the "dots" set first.

3.2.4.1. The "red dots" sub-experiment

We used linear mixed models entering the "number of red dots" as the only fixed factor, and subjects and items as random factors with intercepts, entering also color as a random slope in items. Our models revealed a strong main effect of the "number of red dots" factor ($p < 0.0001$, $\chi^2(1)=22.926$), with the mean acceptability score of 5.08±0.35 (here and below, variation is reported as a standard error) pertaining to sequences with 2 red dots, as opposed to the mean score of 1.74±0.28 pertaining to sequences with 1 red dot. The results are illustrated on Figure 5.

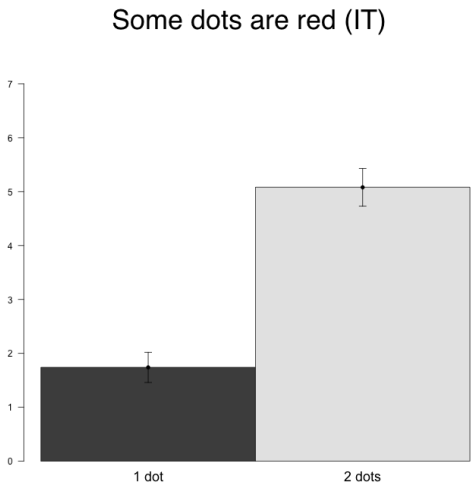


Figure 5. Mean acceptability scores (standard error) on the number of red dots in sequences of 3, 5 and 7 dots, Experiment 1

3.2.4.2. The "ladybugs" sub-experiment

In analyzing the results of this sub-experiment, we separately analyzed the speakers' performance on the conditions 1-4 pertaining to the sentence (32d) and the performance on conditions 5-8 pertaining to the sentence (33d).

3.2.4.2.1. Performance on conditions 1-4

We used linear mixed models with "single-dottedness" as the only fixed factor, entering subjects and items as random factors with intercepts. Our models revealed a strong main effect of the "single-dottedness" factor ($p < 0.0001$, $\chi^2(3)=24.94$), implying that the number of single-dotted ladybugs significantly affects the acceptability score on the sentence in (32d). The mean score on condition 1 (=5 single-dotted ladybugs out of 5) was 2.01±0.31, on condition 2 (=3 out of 5): 3.19±0.39, on condition 3 (=1 out of 5): 3.625±0.40, on condition 4 (=0 out of 5): 6.60±0.17. Posthoc analyses in the form of Tukey's multiple pair comparisons revealed a significant difference in the mean scores between every pair of conditions ($p < 0.001$), except for the pair condition 2 - condition 3 (p =0.27, z = 1.795) where no significance difference was observed. The results are illustrated on Figure 6.

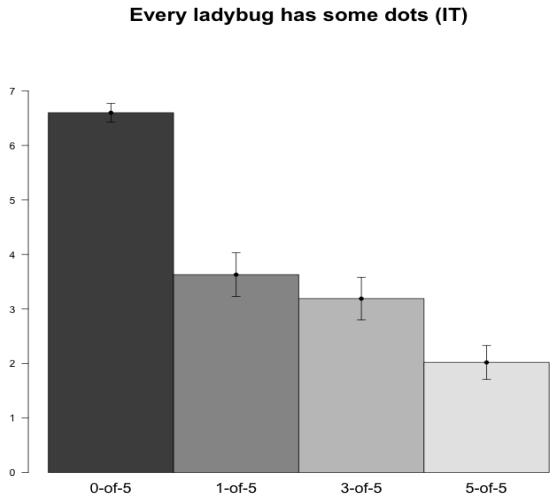


Figure 6. Mean acceptability scores (standard error) on the number of single-dotted ladybugs with respect to (32d), Experiment 1

3.2.4.2.2. Performance on conditions 5-8

Linear mixed modes constructed similarly to the previous case, revealed a strong main effect of the "single-dottedness" factor ($p < 0.0001$, $\chi^2(3)=24.72$), implying that the number of single-dotted ladybugs significantly affects the acceptability score on the sentence in (33d). The mean score on condition 5 (=3 single-dotted ladybugs out of 3) was 2.21±0.32, on condition 6 (=2 out of 3): 3.37±0.40, on condition 7 (=1 out of 3): 3.80±0.37, on condition 8 (0 out of 8): 6.24±0.23. Posthoc analyses in the form of Tukey's multiple pair comparisons revealed a significant difference in the mean scores between every pair of conditions ($p < 0.001$), except for the pair condition 6 - condition 7 (p =0.51, z = 1.368) where no significance difference was observed. Figure 7 illustrates the results.

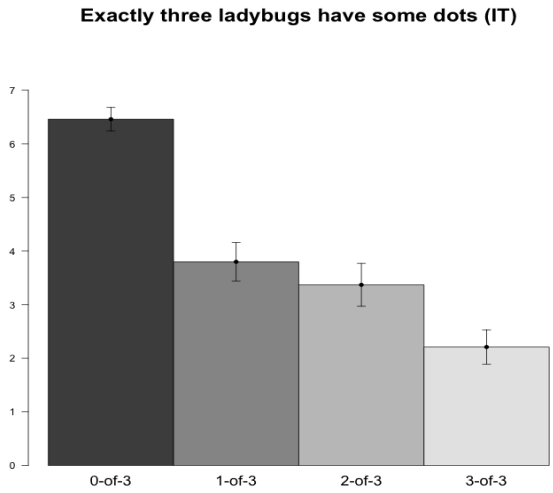


Figure 7. Mean acceptability scores (standard error) on the number of single-dotted ladybugs with respect to (33d), Experiment 1

3.2.5. Discussion

The results obtained in the "red-dots" sub-experiment revealed a strong tendency against accepting (30) as an appropriate description of contexts which contain singular red objects. We interpret these results to support the prediction made by both theories discussed above that (30) carries a plurality inference. This conclusion is important also in view of the fact that our materials do not contain bare plurals but weak existential quantifiers. As we discuss that below, the choice was motivated by the rarity of bare plurals in the language we tested.

The results from the "ladybugs" sub-experiment suggest the following conclusions:

(i) Both theories correctly predict that (32d) is significantly less appropriate when the property of having singular dots is distributed among all ladybugs. The difference in acceptability between Condition 1 and Condition 2 reveals that any amount of multiply dotted ladybugs makes (32d) more appropriate. This outcome supports the background assumptions of both theories according to

which the plural nominal morphology is consistent with atomic, as well, as non-atomic interpretations. If that was not the case, one should expect that (32d) would be equally inappropriate in contexts presented in Conditions 1, 2, and 3.

(ii) The evidenced significant difference in accepting (32d) as appropriate in the context in Condition 3 and Condition 4 goes against the predictions of the *Maximize Presupposition* theory. At the same time, the higher-order implicature theory also predicts no difference between these two conditions on one of the readings calculated for (32d). It is important to acknowledge, however, that the difference is expected under the postulated pragmatic ambiguity: under its strong reading (32d) should be acceptable only if each ladybug is multiply dotted. However, the results would give definite support for Spector's view only if they could show that participants had a strong preference for the strong reading of the plural. But if that was the case, no difference, or perhaps a marginally observed one (due to the hypothetically less preferred reading) should have been observed between Conditions 1 and 2, contrary to fact.

(iii) The observed difference between Conditions 7 and 8 is predicted by the higher-order implicature theory. However, the results go against the predictions of the *Maximize Presuppositions* theory. At the same time, under Spector's theory, in contrast to the *Maximize Presupposition* theory, no difference is expected between Conditions 5 and 6.

3.3. Experiment 2

Both theories considered above were based on English bare plural data. One might argue that the results of Experiment 1 which we took to support partially both the higher-order implicature theory and the *Maximise Presuppositions* theory could be confounded by the use of the narrow scope existential plural quantifier *alcuni* NP in Italian instead of bare plurals. However, the distribution of bare plurals in Romance is extremely limited or non-existent (cf. Chierchia 1998, Zamparelli 2002, Storto 2003, among others). Italian compensates the lack of bare plurals by using either a weak existentially quantified NPs as above, or a partitive-like construction *dei* NP. We, therefore, designed a second experiment using the same target stimuli as in the first experiment but involving partitive NPs in the relevant position. The goal of Experiment 2 was to see if we can replicate the results from Experiment 1. The predictions of the two theories spelled out in the beginning of Section 3.1. are the same in this case.

3.3.1. Design and materials

Experiment 2 was based on the contexts used in the ladybugs set from the first experiment but contained sentences with plural nominals without the determiner *some*. Still we were not able to construct Italian stimuli with bare plurals because these are not acceptable in the relevant contexts (see above). We used the same 4 target conditions as conditions 1-4 in Experiment 1 (see Section 3.1.1), but this time each of those was accompanied by a different sentence that had a plural nominal expression in the scope of a universal quantifier as in (35):

(35) Ogni coccinella ha dei puntini.

´Every ladybug has dots.´

We also used 4 other conditions identical to conditions 5-8 in Experiment 1, constructed with a plural nominal in the scope of the non-monotone quantifier

exactly three. This time, again, the plural nominal phrase in the nuclear scope of the quantifier avoided the use of narrow-scope *some*, as shown in (36).

(36) Esattamente tre coccinelle hanno dei puntini.

´Exactly three ladybugs have dots.´

Similarly to Experiment 1, there were two items per each condition, so that the total of 16 target items were constructed.

Additionally we had 4 items that served as controls. They all included an unembedded plural nominal. Of these, two items included a picture with a ladybug that had 2 dots. Yet another pair of items pictured a ladybug with no dots at all. The corresponding pictures were accompanied by the sentence in (37):

(37) Questa coccinella ha dei puntini.

´This ladybug has dots.´

3.3.2. Participants

36 self-reported monolingual native Italian-speaking adults participated in this study (20 female). They were recruited through personal contacts in the area of Udine, Italy. None of the participants in Experiment 2 had participated in Experiment 1. The participants were aged between 26-42 years (mean age: 32, standard deviation: 3,7). The experiment was again preceded by a short online questionnaire regarding participant´s age, vision and language background. None of the participants reported color-blindedness.

3.3.3. Procedure

The procedure was identical to that in Experiment 1. Each subject received all conditions and controls but presentation was fully randomized per participant within each set.

3.3.4. Results

We excluded data from 6 participants who correctly evaluated less than 79% of the control items. This left 30 participants for further analysis.

Similarly to Experiment 1, we analyzed the groups of conditions 1-4 and 5-8 separately, with respect to the acceptability score in sentences (35) and (36), respectively.

3.3.4.1. Performance on conditions 1-4

Again, linear mixed models with "single-dottedness" as a main factor and subjects and items as random factors with intercepts were constructed. We again found a robust main effect of "single-dottedness" ($p < 0.0001$, $\chi^2(3)=24.45$), implying that the number of single-dotted ladybugs significantly affects the acceptability score on the sentence in (35). The mean score on condition 1 was 2.70±0.34, on condition 2: 4.08±0.37, on condition 3: 4.68±0.40, on condition 4: 6.90±0.05. Posthoc analyses in the form of Tukey's multiple pair comparisons revealed a significant difference in the mean scores between every pair of conditions ($p < 0.001$), except for the pair condition 2 - condition 3, where the mean difference was only marginal ($p =0.07$, $z =2.436$). The results are illustrated on Figure 8.

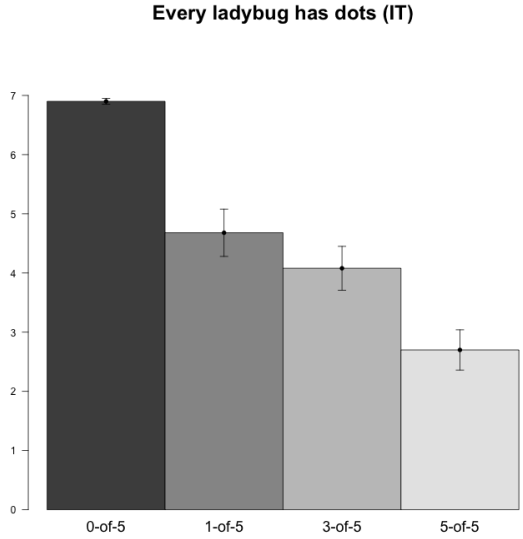


Figure 8. Mean acceptability scores (standard error) on the number of single-dotted ladybugs with respect to (35), Experiment 2

3.3.4.2. Performance on conditions 5-8

We found a robust main effect of "single-dottedness" ($p < 0.0001$, $\chi^2(3)=23.643$), implying that the number of single-dotted ladybugs significantly affects the acceptability score on the sentence in (36). The mean score on condition 5 was 3.30±0.36, on condition 6: 4.33±0.36, on condition 7: 4.73±0.39, on condition 8: 6.83±0.08. Posthoc analyses in the form of Tukey's multiple pair comparisons revealed a significant difference in the mean scores between every pair of conditions ($p < 0.001$), except for the pair condition 6 (2 single-dotted ladybugs out of 3) - condition 7 (1 single-dotted ladybug out of 3), where the mean difference was insignificant ($p =0.35$, $z = 2.436$). The results are illustrated on Figure 9.

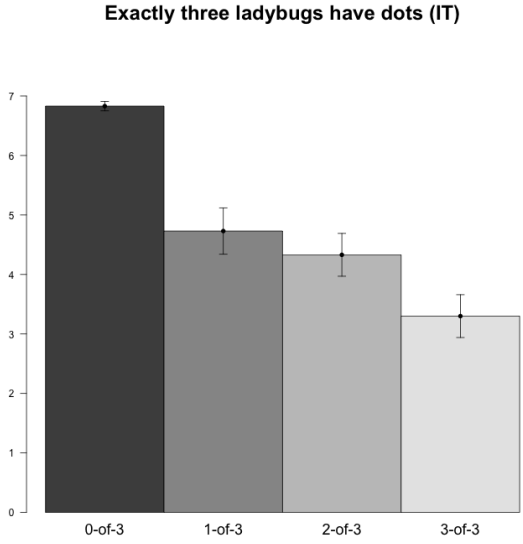


Figure 9. Mean acceptability scores (standard error) on the number of single-dotted ladybugs with respect to (36), Experiment 2

3.3.5. Discussion

The results from the second experiment replicated those obtained in Experiment 1. The first conclusion we make is that we observe the same patterns of significant differences between the pairs of conditions 1 and 2; 3 and 4; 5 and 6; 7 and 8, respectively. We also observe that the partitive construction that was used in Experiment 2 triggered higher mean acceptability scores on every condition. We then conclude that we provide experimental evidence that supports the pragmatic approach to nominal plural interpretation given the observed differences between conditions 1 and 2, and 5 and 6, respectively. The rest of the pairs of conditions that were significantly differently evaluated repeated the pattern of Experiment 1.

4. General discussion and conclusions

Recall that our main goal in this study was to distinguish two current theories of plural interpretations. The results we received support partially each of these two theories. We believe that the most important conclusion we can draw from both experiments is that plural nominal interpretations can contain both atomic and non-

atomic elements. In all four relevant comparisons between pairs of conditions where one of the conditions is viewing exclusively single-dotted ladybugs we observed significant difference in evaluating sentence appropriateness. We interpret it as support for the pragmatic approach to plural nominal morphology against the semantic approach to it. The latter predicts no such distinction, and no exceptions in distributing the property of having multiple dots among each quantified individual. It is also worth mentioning that Conditions 1 and 5 in both experiments, which received the lowest mean values, have contexts that verify the literal meaning of the target sentences according to the pragmatic theory. Consequently, they were not ruled out as definitely inappropriate.

At the same time, we must conclude that each of the two pragmatic theories in our focus received only partial support but we were able to identify the empirical domains that verified each theory. Interestingly, the patterns of sentence evaluation in different contexts were very similar and did not show dependence on the quantifier embedding the target construction. This is not to be expected under Spector's theory. Since within that theory, the universally quantified (32d) and (35) are claimed to be ambiguous in contrast to (33d) and (36), we expected that ambiguity to result in different relative judgment patterns. Crucially, in the case of the first but not the latter pair of sentences the significant difference between contexts with only singly dotted ladybugs (Conditions 1 and 5) and contexts with some but not all of the ladybugs being multiply dotted (Conditions 2,3 and 6,7), is expected (cf. Figure 1 and Figure 4) since the weak readings that are available only for (32d) and (35) are verified in Conditions 2 and 3. Furthermore, only the part (33d) and (36) but not the pair (32d) and (35) was expected to trigger a difference between contexts with exclusively multiply-dotted ladybugs (Conditions 4 and 8) and contexts with mixed pattern of dottedness (Conditions 2,3 and 6,7). However, while the prediction is confirmed with respect to (33d) and (36) (cf. Figures 7 and 9) it is not supported by the obtained data with respect to (32d) and (35) (cf. Figures 6 and 8).

Turning to Sauerland's theory, we can conclude, in a similar vain, that it found partial experimental support. In this case, the predictions about a contrast between Conditions 1 and 5 on the one hand and Conditions 2, 3 and 6, 7, on the other were matched by both experiments' results as evidenced by Figures 6-9. However, none the differences between multiply-dotted ladybugs (Conditions 4 and 8) and contexts with mixed pattern of dottedness (Conditions 2,3 and 6,7) in the two experiments was expected.

We conclude then that our study favors the pragmatic approaches to plural interpretations and suggests in which particular domains predictions do not find experimental support. In the big picture, we believe that Sauerland's (2007) theory has a slightly better fit. If we are correct in that conclusion then one might further argue that plurality inferences should not be reduced to (second-order) scalar implicatures.

References

Chemla, E. and Spector, B. (2011). "Experimental evidence for embedded scalar implicatures", *Journal of Semantics 28:3*, 359-400.

Chierchia, G. (1998). "Reference to kinds across languages". *Natural Language Semantics 6*, 339-405.

Chierchia, G. (2004) "Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface", in *Structures and Beyond 3*, ed. Belletti, A. (Oxford: Oxford University Press), 39-103.

Chierchia, G. (2013). *Logic in Grammar: Polarity, Free Choice, and Intervention*. Oxford: Oxford University Press.

Chierchia, G., Fox, D., and Spector, B. (2012). "Scalar implicature as a grammatical phenomenon," in *An international handbook of natural language meaning semantics*, eds. P. Portner, C. Maienborn, and K. von Heusinger, Berlin: Mouton de Gruyter, 2297-2331.

Fox, D. (2007). "Free choice and the theory of scalar implicatures," in *Presupposition and implicature in compositional semantics*, eds. U. Sauerland and P. Stateva, Palgrave, 71-120.

Heim, I. (1991). "Artikel und Definitheit," in *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, eds. A. von Stechow, A. and D. Wunderlich, Berlin: de Gruyter, 487–535.

Ivlieva, N. (2013). *Scalar Implicatures and the Grammar of Plurality and Disjunction*, Massachusetts Institute of Technology, PhD dissertation.

Magri, G. (2014). *An account for the homogeneity effects triggered by plural definites and conjunction based on double strengthening*. Ms. CNRS, Univ. Paris 8.

Mayr, C. (2015). "Plural definite NPs presuppose multiplicity via embedded exhaustification," in *Proceedings of Semantics and Linguistic Theory 25*, eds. S. D'Antonio, M. Moroney, and C. Rose Little, 204–224.

Percus, O. (2006). "Antipresuppositions," in *Theoretical and empirical studies of reference and anaphora*, ed. A. Uyema, Japan Society for the Promotion of Science, 52-73.

Sauerland, U. (2003). "A new semantics for number," in *Proceedings of Semantics and Linguistic Theory 13*, eds. R. Young and Y. Zhou, Ithaca, NY: Cornell University, 258-275.

Sauerland, U. (2008). "Implicated presuppositions," in *Sentence and context: Language, context, and cognition*, ed. A. Steube, Berlin: Mouton de Gruyter, 581-600.

Sauerland, U., Andersen, J., and Yatsushiro, K. (2005). "The plural is semantically unmarked," in *Linguistic evidence: Empirical, theoretical, and computational perspectives*, eds. S.Kepser and M. Reis, Berlin: Mouton de Gruyter, 413-434.

Schlenker, P. (2012). "*Maximize Presupposition* and Gricean reasoning", *Natural Language Semantics 20*, 391-429.

Singh, R. (2011). "*Maximize Presupposition!* and local contexts", *Natural Language Semantics 19 (2)*, 149-169.

Spector, B. (2007). "Aspects of the pragmatics of plural morphology: On higher-order implicatures," in *Presupposition and implicature in compositional semantics*, eds. U. Sauerland and P. Stateva, Palgrave, 243-281.

Storto, G. (2003). "On the status of the partitive determiner in Italian," in *Romance Languages and Linguistic Theory 2001: Selected Papers from Going Romance 2001*,eds. J. Quer, J. Schroten, M. Scorretti, P. Sleeman, and E. Verheugd , Amsterdam: John Benjamins, 315-330.

Tieu, L., Cory, B., Romoli, J., and Crain, S. (2014). "Plurality inferences are scalar implicatures: Evidence from acquisition, " in *Proceedings of Semantics and Linguistic Theory 24*, eds. J. Grieser, T. Snider, S. D'Antonio, and M. Wiegand, Portland, USA: Linguistic Society of America, 122-136.

Zamparelli, R. (2002). "Definite and bare kind-denoting noun phrases," *in Romance Languages and Linguistic Theory 2000: Selected Papers from Going Romance 2000*, eds. C. Beyssade, R. Bok-Bennema, F. Drijkoningen, P. Monachesi , Amsterdam: John Benjamins, 305-344.