

## Typicality made familiar: A commentary on Geurts & van Tiel (2013)

Chris Cummins  
*University of Edinburgh*

Submitted 2013-11-04 / First decision 2013-11-25 / Revision received 2013-11-29 /  
Second decision 2013-12-11 / Revision received 2013-12-16 / Accepted 2013-12-31 /  
Final version received 2013-12-31 / Published 2014-06-03

### Abstract

In their recent paper, Geurts & van Tiel (2013) review a range of evidence on the availability of embedded scalar enrichments (upper-bound construals, or UBCs). They argue that these readings are not readily available, except when triggered by contrast effects, and conclude that the experimental data do not support a conventionalist view of implicature. They also consider how some of these data can be analysed as exhibiting typicality effects. In this commentary, I focus on the claim that typicality effects apply to quantifiers, and consider some of its implications for our view of semantics and pragmatics. In particular, I look at whether these effects are general to embedded and non-embedded contexts, whether and how typicality relates to truth-conditional narrowing, and the implications of this view for the nature of pragmatic enrichment. I conclude that typicality effects are indeed in evidence in the data elicited so far, and that this opens up several promising new avenues for the study of quantification in natural language, as well as challenging our interpretation of existing data.

**Keywords:** implicature, typicality, quantification

### 1 Introduction

A flurry of recent experimental research has addressed the vexed question of so-called “embedded implicatures,” or local upper-bound construals (UBCs): cases where the pragmatic enrichment of a weak scalar term seems to take place locally, under the scope of an operator. For instance, it is claimed that

©2014 Chris Cummins

This is an open-access article distributed under the terms of a Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>).

(1) can give rise to the reading (2), as a consequence of which it conveys that no square is connected to all of the circles.

- (1) Every square is connected to some of the circles.
- (2) Every square is connected to some but not all of the circles.

Geurts & van Tiel (2013) review three experimental papers on this topic: Geurts & Pouscoulous 2009, which argued that embedded UBCs are very infrequently in evidence, and Clifton & Dube 2010 and Chemla & Spector 2011, which argued that embedded UBCs are relatively widespread in a way that suggests the inadequacy of a Gricean pragmatic account.<sup>1</sup> Geurts & van Tiel criticise these latter two papers on two distinct grounds. First, they argue that the materials used give rise to contrast effects, which suggests that the results can be parsimoniously accounted for on pragmatic grounds after all; and secondly, they argue that the results can be even better accounted for in terms of typicality.

On the former point, Geurts & van Tiel emphasise what they see as an under-appreciated unity between the competing theoretical approaches. It is, in their view, common to all approaches that truth-conditional narrowing (for instance, “some” being understood to mean more specifically “some but not all”) can take place in contrastive environments, just as it is assumed by all approaches that some form of conversational implicature must exist in the system (for instance, to account for uncertainty implicatures). What is disputed is essentially just how truth-conditional narrowing arises, although this goes on to have substantial implications for when it can occur.

In support of the claim that typicality offers a better explanation of the experimental data, Geurts & van Tiel (2013) adopt a typicality measure introduced by van Tiel (2013), which can be used to quantify how typical a situation is as an instantiation of a sentence. They demonstrate how this method can be applied to the materials used by Chemla & Spector (2011). For the sentence (1), this would involve calculating the typicality of the description “connected to some of the circles” for each square separately, and combining these typicality ratings into a single measure, by taking either their arithmetic or harmonic mean. Geurts & van Tiel 2013 show that a measure of this type

---

<sup>1</sup> To infer (2) from the utterance of (1) in a Gricean fashion, we would require some auxiliary assumption, for instance that every square is connected to the same number of circles. It seems possible, even within a classically Gricean analysis, that an utterance such as (1) might influence our beliefs as to whether an appropriate auxiliary assumption also holds. However, further speculation on this topic is beyond the scope of this paper.

accurately predicts the ratings obtained by [Chemla & Spector](#), and thus that the results can be explained efficiently by appeal to typicality effects and without reference to embedded implicatures. Consequently, [Geurts & van Tiel](#) conclude that [Chemla & Spector 2011](#) does not furnish any novel experimental support for the conventionalist analysis of scalar implicature, and thus argue in favour of a pragmatic account.

[Geurts & van Tiel's](#) argument with respect to the recent experimental data thus travels some distance over the course of the paper. They first set out to establish that embedded UBCs are compatible with all the currently competing theoretical proposals, and then describe how contrast effects could be invoked as part of an explanation for the findings of both [Clifton & Dube 2010](#) and [Chemla & Spector 2011](#). However, for both data sets, they go on to argue that a typicality-based account is more satisfactory, and that there is no clear evidence for embedded UBCs in either experiment.

I concur with [Geurts & van Tiel](#) that the data elicited so far fail to constitute a compelling case for conventionalism, and have little to add to their criticisms of the corresponding analyses. However, their typicality-driven account appears to raise a number of issues with potentially broad consequences for semantic and pragmatic theory and experimentation. For instance, given the potential availability of typicality-based explanations alongside implicature-based accounts, exactly how widespread are such typicality effects? Specifically, can many, or most, apparently embedded UBCs actually be ascribed to typicality? Does the typicality effect extend to unembedded UBCs, even in contexts where implicatures are predicted on purely Gricean grounds? Does it matter whether or not implicatures influence truth judgments? And doesn't the postulation of typicality effects tend to blur the distinction between the Gricean and the conventionalist account of implicature?

In this commentary, I attempt to address these points. First, I revisit the existing experimental data, with a view to establishing what participants are actually doing in these experimental paradigms, and thus what precisely we have to explain. Then I consider how and why the typicality effects associated with quantifiers might vary between embedded and unembedded contexts of usage. I look briefly at the contrast between typicality and defaultist views of implicature, and consider the place of typicality in a pragmatic system. Finally, I consider how it might be possible to derive novel predictions about implicature if typicality is assumed.

## 2 Typicality versus (embedded) UBCs

All three papers discussed by Geurts & van Tiel (2013) report experiments using broadly similar materials: these were ostensibly meaningless geometrical diagrams consisting of connected squares, circles, and letters. Participants were then asked to rate sentences such as (1), repeated below, as descriptions of these diagrams.

- (1) Every square is connected to some of the circles.

The use of such materials has triggered a certain amount of mutual criticism and scepticism between the various sets of authors as to the validity of the experiments. Objections raised over the course of this debate include whether the diagrams are comprehensible to participants, in terms of making the sentences verifiable; and indeed whether the configuration of the diagrams suggests that some propositions are likely to be more relevant than others.

However, perhaps the most important issue concerns what participants consider themselves to be rating, when they respond to these materials. In part of Clifton & Dube's (2010) experiment, participants were presented with sentences similar to (1) and given a choice between two diagrams, one of which is true given the embedded UBCs (i.e., each square was connected to some but not all of the circles) and one of which is false under this interpretation (i.e., one square was connected to all of the circles, while the others were connected to some but not all of the circles). The participants were asked which diagram was best described by the sentence, but were given a four-way choice - they could select either diagram, respond "both" or respond "neither". Among the participants, those who expressed a preference broke strongly in favour of the diagram that was true given the UBC (39%-3%, with 57% responding "both").

The pragmatics of the four-way choice is itself very complicated. (Is it even felicitous to say that both of the diagrams were best? If not, what question were the 57% of respondents who said "both" actually answering?) The rationale for offering such a choice is, fairly clearly, not to force participants to state a preference if they do not really have one. However, even granted that justification, there is no evidence that the preference (sometimes) exhibited reflects any kind of truth-conditional effect. The only certain conclusion is that the 57% of participants who chose "both" (and the 1% who chose "neither") were not materially affected by any kind of truth-conditional narrowing

in their choice of response, and nor were the 3% who chose the UBC-violating diagram. There is no clear reason to further suppose that the remaining 39% were actually motivated by truth-conditional concerns in their statement of preference.

For [Chemla & Spector \(2011\)](#), the story is very similar. Their participants were asked to indicate on a continuous scale how “true” or “appropriate” sentences were as a description of diagrams. Again, diagrams that were coherent with the UBC of weak embedded scalar “some” were rated higher than those which were not. But again, the results are not clearly reflective of truth-conditional effects: specifically, the difference in rating does not appear to be due to some participants rating the UBC-violating diagrams as clearly false, as would be expected on a truth-driven account.

[Geurts & van Tiel \(2013\)](#) articulate what I think is a more constructive objection to the authors’ analyses of their respective data sets, by demonstrating that the results can be modelled by appeal to typicality. Importantly, this account also handles the data that arises from other non-critical conditions of [Chemla & Spector’s](#) experiment: it provides an explanation for how one clearly false sentence comes to be rated as more “appropriate” (presumably not “true”) than another. This observation seems to call for a typicality-based explanation, and the unified account put forward by [Geurts & van Tiel \(2013\)](#) seems to offer an especially parsimonious explanation of the full data set.

More generally, [Geurts & van Tiel](#) demonstrate that the patterns in the data under discussion admit (at least) two distinct interpretations: as local UBCs (whether or not these are attributable to contrast effects) or as typicality effects. This ambiguity of interpretation arises in part because the instructions to participants in the above-mentioned papers vacillated between reference to truth judgments and acceptability judgments. Suppose, as in the canonical examples, that we have a situation in which a sentence with “all” would be true, and we test the usage of the corresponding sentence with “some”. If “some” attracts a UBC, it should be judged as false in a truth-value judgment task, and also receive a low rating in an acceptability judgment task. If “some” does not attract a UBC but is subject to typicality effects, it should be judged as true in a truth-value judgment task, but receive a low rating in an acceptability judgment task. Consequently, only the truth-value judgment task should be able to distinguish between the two possible accounts.

In defence of the use of acceptability judgments, we might argue that asking whether a sentence is a good description of a situation is more natural than asking whether, in some technical sense, it is true or false. Moreover, as experimenters, we may suspect that typical language users, with no specialist training in logic, will in any case interpret questions about truth or falsity as questions about acceptability or naturalness. That being the case, these tasks may not be diagnostic of UBCs as such. The viability of typicality-based explanations for embedded implicature data thus raises the spectre of similar explanations being brought to bear on a range of empirical data on implicature in unembedded contexts. Such an approach would cohere with the findings of [Katsos & Bishop \(2011\)](#): they showed that hearers unanimously penalise underinformative utterances (“some” used to describe a situation in which “all” is true) in acceptability judgment tasks, but only some hearers judge such utterances false. Hence, [Katsos & Bishop’s \(2011\)](#) findings also challenge the assumption that dissatisfaction with a quantity expression is a good indicator of the availability of a scalar implicature that would render that expression false. Taken together with this work, [Geurts & van Tiel’s \(2013\)](#) analysis suggests that we may need to exercise considerable caution in interpreting apparent preferences for UBCs as evidence for implicature in any technically defensible sense of the word.

### 3 Typicality in one and two dimensions

In their discussion of [Clifton & Dube’s \(2010\)](#) experiment, [Geurts & van Tiel](#) show that a similar result can be obtained if the experiment is adapted to use images of entities that are uncontroversially typical and atypical category members. Their version of the experiment presented participants a picture of a robin and an ostrich, and asked which picture is best described by “This is a bird”, again offering a four-way choice (picture A, picture B, both, neither). Again, a sizeable minority of participants states a preference for one picture, and of the participants who state a preference, most choose the robin rather than the ostrich. Here, there is little room to entertain the possibility that participants are responding on the basis of truth or falsity: “This is a bird” is just as true of the ostrich, definitionally, but participants nevertheless feel that it better describes a robin.

Having demonstrated this analogy, [Geurts & van Tiel](#) offer a more concrete proposal for the “some” case. They argue that “some” exhibits typicality structure, with the most typical meaning of “some (of)” being situated around

the mid-point of the scale, i.e., referring to about half of the items under discussion. “Every” on this account also exhibits some typicality structure, the typical meaning being the literal meaning (“all (of)”).

Evidence in favour of the claimed typicality effect comes from two directions. First, [van Tiel 2013](#) reports an experiment in which participants are asked to rate the acceptability of two sentences, “Every circle is black” and “Some of the circles are black”, used as descriptions of a range of situations in which 10 circles are presented and some number of them are black. For the former sentence, only the condition with 10 black circles achieved a high rating, but ratings monotonically increased as the number of black circles increased. For the latter, conditions with 2 to 9 black circles received high ratings, peaking at 5; the condition with 10 black circles received a lower rating, and the 1 and 0 black circle conditions lower still. Secondly, when these ratings are used as the basis for a typicality measure, and this measure is used to predict the acceptability of the sentences used in [Chemla & Spector’s \(2011\)](#) experiment, the result fits well with their data.

One possible objection to the typicality structure elicited by [van Tiel \(2013\)](#) is that the task he uses presents quantifiers in an unembedded context. Most strikingly, it seems possible that the preferences thereby elicited are themselves influenced by implicature. If a participant establishes a UBC for the sentence “Some of the circles are black” — which they could do, uncontroversially, by scalar implicature — then clearly we would expect them to rate the sentence as suboptimal as a description of a situation in which all of the circles are black. It is noteworthy that the mean rating for “Some of the circles are black” in the 10 black circles condition is lower than the mean rating for “Every circle is black” in the 9 black circles condition, a sentence that we would not hesitate to call “false”. So, on the face of it, the claim that “all” is a highly non-typical case of “some” appears to be bound up with the availability of the scalar implicature. By (Gricean) hypothesis, such an implicature is restricted to the unembedded case, and should not be able to tell us anything about the behaviour of “some” in embedded contexts.

In my view, the above objection is not a particularly serious one, in that there is convergent evidence for approximately the kind of typicality structure that [Geurts & van Tiel](#) posit for “some” which cannot be explained away as a consequence of implicature. [Bååth, Sauerland & Sikstrøm \(2010\)](#) presented data from an elicitation study in which participants were presented with a display of 432 dots, some of which were blue and some yellow (in varying proportions), and were asked to say how many of the dots were yellow. They

elicited a wide range of quantifying expressions, including “few”, “some”, “half”, “many”, “most”, and “almost all”. Notably, the use of “some” was largely restricted to the middle of the range, being used very infrequently for small numbers of dots and not at all in situations where the proportion of yellow dots neared 100%. If we can assume that our interpretative preferences are attuned reasonably well to other speakers’ production preferences, we can further assume on this basis that “some” should be preferentially interpreted as referring to values in the middle of the available range. It seems clear that [Bååth, Sauerland & Sikstrøm’s](#) data are not wholly explicable just in terms of implicatures, as such an account would predict that “some” should become unavailable at the point at which “all” (or conceivably “most”) becomes true. Instead, in their data, “some” fades out gradually as the proportion increases, becoming unattested at a point some way short of “all”, as more informative descriptions (“most”, “many”, “almost all”, etc.) come to be preferred. This pattern appears to conform to expectations about how typicality effects should behave in the domain of quantifiers.

If the typicality effects elicited by [van Tiel \(2013\)](#) are not due to implicature, that removes the major objection in principle to the application of their results to the embedded case, which is what [Geurts & van Tiel \(2013\)](#) do in their discussion of [Chemla & Spector 2011](#). However, we might still ask whether the landscape of quantification is really the same in the embedded case as in the unembedded case – that is to say, whether the possible (preferred) choices of expression divide up the set of possible situations in the same way as they were shown to do in [Bååth, Sauerland & Sikstrøm’s \(2010\)](#) study. This issue appears to have received very little attention thus far.

Abstractly, an unembedded scalar term picks out a region in one-dimensional space (that is, part of a line, which in the case of partitive quantification runs from “none” to “all”), and the posited typicality effect identifies some sub-part of that region as typically containing the denotation of the scalar term. By contrast, an embedded scalar term picks out a region in two-dimensional space (that is, part of a plane, which in the case of partitive quantification extends from “none...any” to “all...all”). In this case, the nature or shape of the typicality effects is not well understood. We can visualise this if we look at a two-dimensional data array that plots a set of individuals against a set of properties and simply indicates whether or not each individual has the corresponding property.

To take a specific example, consider the situation in [Figure 1](#), which depicts (the imagined) marks for a group of students over the 10 questions in

---

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Student 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Student 2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Student 3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Student 4	✓	✓	✓	✓	✓	✓	✓	×	×	×
Student 5	✓	✓	✓	✓	✓	✓	✓	×	×	×
Student 6	✓	✓	✓	✓	✓	✓	×	×	×	×
Student 7	✓	✓	✓	✓	×	×	×	×	×	×
Student 8	✓	✓	✓	✓	×	×	×	×	×	×
Student 9	✓	✓	✓	×	×	×	×	×	×	×
Student 10	✓	✓	✓	×	×	×	×	×	×	×

**Figure 1** Example of two-dimensional data array

---

an exam. These data could be summarised with sentences (3) or (4), whereas (5) would be false. As before, the utterance of (3) or (4) uncontroversially implicates (under the usual conditions for implicature) the falsity of (5).

- (3) Some of the students got all of the questions right.
- (4) All of the students got questions 1-3 right.
- (5) All of the students got all of the questions right.

A more controversial sentence, as far as the UBC is concerned, would be (6).

- (6) All of the students got some of the questions right.

There is an additional ambiguity here: (6) could convey that all of the students got certain specific questions right, as in (4). But I will ignore this reading for the moment. On the other reading, it conveys that there exists no student who answered no questions correctly. With the UBC, it would also convey that there existed no student who answered all the questions correctly: in this case it would be false for the data presented in Figure 1.

Now, if we apply the typicality findings for “some” to this embedded case, we would expect to find that the data depicted in Figure 1 can be construed as an atypical case of (6). “Some” (out of 10 items) typically means 2-9 out of 10, so the description “got some of the questions right” with a typical construal applies to only 7 of the 10 students in Figure 1, not to all of them. That is to say, (6) is true of this situation, but only if we grant that the quantifier

“some” in (6) is conveying a somewhat atypical meaning. There appears to be an element of truth to this claim: pilot data from [Cummins 2012](#) showed that participants in an elicitation task preferred to use option (3), whereas (6) was the preferred option if none of the students got all of the questions right — a situation in which the description “got some of the questions right” is true under a typical construal for all 10 of the students (and in which (3) would, incidentally, be false).

However, there is also a sense in which the above claim seems to be too strong. The problem is that the data in Figure 1 are perfectly commonplace, but there is no efficient way to describe them. It is not at all obvious, on inspection, whether they are closer to the prototype of (3) than they are to the prototype of (6). To resolve that question, we would need to understand what kind of expectations speakers have about the data that are not described either by assertion or implicature, when they encounter a description like (3). Does (3) convey any sense of how the other students — those falling outside the set delimited by “some” — performed in the exam? Yes: it certainly implicates that they didn’t get all the questions right. But it also suggests that they didn’t get all the questions wrong, as otherwise (7) would be a competing utterance and might conceivably be more relevant. Whether (3) conveys anything else about the distribution is an empirical question.

(7) Some/most of the students got all of the questions wrong.

In short, the landscape of possibilities for the use of embedded quantifiers is a much more complex one than that for simple quantifiers. And [Bååth, Sauerland & Sikstrøm \(2010\)](#) show that the situation is already fairly complex in the latter case, if we admit the existence of typicality effects and are interested in how quantifying expressions are actually used, rather than just how they should be treated truth-conditionally. Given the complexity of the situation, I feel obliged to treat the analysis of [Geurts & van Tiel \(2013\)](#) with a certain degree of caution, in that it supposes that the behaviour of simple quantifiers generalises straightforwardly to the embedded case, which may not be tenable. But more generally, this type of analysis suggests that the experimental data are running way ahead of any really satisfactory or comprehensive explanation. It would be unhelpful, and untrue, for me to assert that so much is going on in the paradigms of [Geurts & Pouscoulous \(2009\)](#), [Clifton & Dube \(2010\)](#), and [Chemla & Spector \(2011\)](#) that we can’t yet make any kind of sense of the results. However, I think it is fair to say that the original analyses of these data sets rely on simplifying assumptions that

may become increasingly difficult to defend, as we explore the behaviour of embedded quantifiers more thoroughly.

#### 4 Typicality and truth-conditional narrowing

A potentially important matter arising from Geurts & van Tiel's (2013) analysis is the truth-conditional status of quantifiers that exhibit typicality effects. They emphasise that typicality effects can arise in domains where there is no difficulty whatsoever as to truth-conditions: there is such a thing as a prototypical even number, or prime number, etc.<sup>2</sup> Moreover, they appear to be drawing a more or less clear distinction between typicality effects and truth-conditional narrowing: although the latter can proceed in accordance with pragmatic principles, and by doing so can explain the experimental data that they discuss, ultimately they set aside this analysis in favour of a distinct account based upon typicality effects that does not involve truth-conditional narrowing or the generation of UBCs.

At the same time, there are reasons to suppose that the distinction between the two approaches might not be quite so clear-cut. Geurts & van Tiel (2013, p. 7) are unambiguously discussing the pragmatic view of truth-conditional narrowing when they make the point that “occasionally the occurrent meaning of a word will be more specific than its dictionary meaning. In some cases, this narrowing may proceed quite smoothly, because world knowledge alone suffices to steer the hearer towards a specific interpretation.” They go on to discuss cases in which contrastive stress is required in order for truth-conditional narrowing to proceed pragmatically, which they argue are not amenable to analysis in conventionalist terms. However, the kind of typicality effect that they posit for “some” appears to fall squarely within their characterisation of truth-conditional narrowing, merely of the kind that proceeds smoothly without the support of contrastive stress. We understand that “some” conveys something other than the purely existential meaning that semantics would suggest, but that this is not merely attributable to a clean scalar implicature derived by contrast with “all” — instead, it reflects an awareness on our part as to the way in which this word is typically used.

The difference between the typicality account and a lexicalist conventionalist account is, as a consequence, rather subtle. On a lexicalist account, we say that “some” means “some but not all”, but that the latter part of this meaning is cancellable. On a typicality account, “some” means some-

---

<sup>2</sup> For example, 2 is both a highly typical even number and a highly atypical prime number.

thing slightly different from this, in that its meaning happens to exclude or marginalise the possibility of “all”. The status of this meaning, like its precise content, is rather difficult to pin down: it appears to be something like a default inference, in the sense of [Levinson 2000](#). However, it is not obviously cancellable — it is not clear what it would mean to be able to cancel a typicality effect. And it is not clear that we can usefully appeal to truth-conditions in trying to characterise such an effect: less typical examples of “some” are just as true as more typical examples (if we exclude the “all” case).

What appears to be on the table here is a form of content that typically results in a narrowing of the range of expected meanings, given an utterance, rather than a narrowing of the range of truth-conditionally admissible meanings. In principle, this is an intriguing prospect, and appears to mesh with some points hinted at by [Levinson \(2000\)](#) about intermediate levels of meaning. It remains to be seen whether such an account can be made rigorous enough to be useful, but if it can, it might be relevant to the interpretation of a great deal of experimental data (which has often been elicited in rather non-naturalistic conditions).

## 5 Typicality-driven (and cancelled) implicatures

It seems clear that [Geurts & van Tiel \(2013\)](#) consider that typicality structure exists alongside the familiar Gricean mechanisms for quantity implicature. They wish to emphasise from the outset that so-called “embedded implicatures” are not implicatures at all, inasmuch as implicatures are only calculated at the utterance level: “it is pointless even to consider the possibility than an implicature might occur in an embedded position” ([Geurts & van Tiel 2013](#), p. 3). By contrast, typicality effects are presumed to be potentially operative in any position.

If typicality effects are widespread, it seems credible that they may influence the availability of certain implicatures. Canonically, the implicature “some” +> “not all” is explicable on the basis that the speaker would have uttered the stronger statement if she were able to do so, and — if we further assume her knowledgeability as to the truth of the stronger statement — we can conclude that the stronger statement is false, as far as the speaker knows. Now, if typicality effects are in play, these seem obliged to influence the admissibility of the speaker’s making stronger statements. One should reason as follows: the speaker made a weaker statement, truth-conditionally. She could have made a stronger statement. However, in doing so, she would have

conveyed not only that the truth-conditions of the stronger statement were satisfied, but also that the situation was a highly typical one in which that statement could be uttered.

In the case of “some” +> “not all”, the relevance of such considerations is at most marginal: although [Geurts & van Tiel \(2013\)](#) hint at the idea that “all” has typicality structure, there is only one situation that satisfies the truth-conditions, so typicality cannot take any effect, on the above analysis. However, if we consider a case like “some” +> “not most”, we may actually be able to derive a novel prediction based on typicality. Specifically, we can observe that “most” semantically seems to mean “more than 50%” but appears typically to convey a larger quantity ([Hackl 2009](#), [Bååth, Sauerland & Sikstrøm 2010](#), [Solt 2011](#)). Take a situation in which 51% of the apples are green. The speaker who utters (8) chooses not to utter (9).

(8) Some of the apples are green.

(9) Most of the apples are green.

We can argue that the choice of (8) need not reflect the falsity of (9), but could instead be taken to reflect the atypicality of the situation given an utterance of (9). For this reason, we should expect the implicature to be unavailable. We reason as follows: we heard the speaker say (8), when they could have said (9). One possible explanation is that (9) is false. Another possible explanation is that (9) is true, but the situation that prevails is not a highly typical instance of this circumstance. All we can conclude with certainty (assuming, as usual, that we are dealing with a cooperative and well-informed speaker) is that the situation that prevails is not a highly typical case of (9).

Therefore, from a typicality standpoint, we should not expect the implicature “some” +> “not most” to be available, but rather the implicature “some” +> “not typical-most”. This is an informationally weaker prediction, in that the negation of “typical-most” entails the negation of “most”, and the precise scope of the prediction naturally depends on what constitutes typical “most”. Nevertheless, the prediction may be worth investigating in support of the relevance of typicality in the quantity domain.

One final point to emphasise is that the relation of typicality effects to intentional communication is not straightforward. Specifically, it could be argued that speakers do not necessarily intend to convey typicality effects: if all possible choices of expression are associated with some kind of typicality structure, then the speaker cannot help but suggest some kind of typicality

effect, irrespective of whether this is appropriate given the speaker's knowledge state. On a classically Gricean account, such as that endorsed by Geurts & van Tiel (2013), this means that the typicality effects associated with the utterance that is actually made are not considered to be communicated. By contrast, the inappropriateness of the typical interpretations associated with alternatives that were not uttered (as discussed above) could potentially be considered to be communicated. Untangling this relationship might be a useful avenue for future research.

## 6 Conclusion

Despite the recent wave of experimental literature on the topic, there is still no consensus as to the nature and prevalence of embedded upper-bound construals. Geurts & van Tiel (2013) draw attention to two major obstacles to the interpretation of the existing experimental data: the presence of contrast effects in the elicitation paradigms used, and the possibility of typicality effects in the interpretation of scalar terms (and other items). They convincingly demonstrate the difficulty of inferring the existence of "embedded implicatures" on the basis of the data available so far. Perhaps more importantly, they provide novel insights into some of the factors that must be taken into account in future attempts to investigate this issue experimentally. In this commentary, I argue that, if typicality effects really are manifest in quantifiers, this has potentially far-reaching consequences for how we should elicit and interpret experimental data in pragmatics, and for how we should construe the nature and interface of semantic and pragmatic meaning.

## References

- Bååth, Rasmus, Uli Sauerland & Sverker Sikstrøm. 2010. Quantifier use in English and German: An online study. In *Workshop on vague quantities and vague quantifiers*. Zentrum für Allgemeine Sprachwissenschaft. Berlin.
- Chemla, Emmanuel & Benjamin Spector. 2011. Experimental evidence for embedded scalar implicatures. *Journal of Semantics* 28. 359-400. <http://dx.doi.org/10.1093/jos/ffq023>.
- Clifton, Charles & Chad Dube. 2010. Embedded implicatures observed: A comment on Geurts and Pouscoulous (2009). *Semantics and Pragmatics* 3(7). 1-13. <http://dx.doi.org/10.3765/sp.3.7>.

- Cummins, Chris. 2012. Using embedded quantifiers. In *Much ado about much: The semantics and pragmatics of quantity expressions*. Universität Konstanz, Konstanz, Germany. [http://www.crcummins.com/Cummins\\_Konstanz.pdf](http://www.crcummins.com/Cummins_Konstanz.pdf).
- Geurts, Bart & Nausicaa Pouscoulous. 2009. Embedded implicatures?!? *Semantics and Pragmatics* 2(4). 1–34. <http://dx.doi.org/10.3765/sp.2.4>.
- Geurts, Bart & Bob van Tiel. 2013. Embedded scalars. *Semantics and Pragmatics* 6(9). 1–37. <http://dx.doi.org/10.3765/sp.6.9>.
- Hackl, Martin. 2009. On the grammar and processing of proportional quantifiers: *Most* versus *more than half*. *Natural Language Semantics* 17. 63–98. <http://dx.doi.org/10.1007/s11050-008-9039-x>.
- Katsos, Napoleon & Dorothy V. M. Bishop. 2011. Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition* 120. 67–81. <http://dx.doi.org/10.1016/j.cognition.2011.02.015>.
- Levinson, Stephen C. 2000. *Presumptive meanings*. Cambridge, MA: MIT Press.
- Solt, Stephanie. 2011. How many mosts? *Sinn und Bedeutung* 15. 565–579. <http://universaar.uni-saarland.de/monographien/volltexte/2011/30/pdf/SinnUndBedeutung15.pdf>.
- van Tiel, Bob. 2013. Embedded scalars and typicality. *Journal of Semantics* 31(2). <http://dx.doi.org/10.1093/jos/ffto02>.

Chris Cummins  
School of Philosophy, Psychology and Language  
Sciences  
University of Edinburgh  
[c.r.cummins@gmail.com](mailto:c.r.cummins@gmail.com)