# Embedded Implicatures Observed: A Comment on Geurts and Pouscoulous (2009),*

**Charles Clifton Jr** and **Chad Dube**
University of Massachusetts Amherst

## Abstract

Conventionalist theories of scalar implicature differ from other accounts in that they predict strengthening of embedded scalar terms. Geurts and Pouscoulous (2009) argue that experimental support for this prediction is largely based on sentence comprehension tasks that inflate the frequency with which terms like *some* are strengthened. Using a picture verification task, they observed no strengthening of embedded scalars. We present data from a multiple-choice picture verification task that is more sensitive to interpretation preferences, and find that readers do show a preference for strengthened interpretations even in embedded phrases. These data cast doubt on Geurts and Pouscoulous's empirical arguments against the existence of embedded implicatures.

Geurts & Pouscoulous (2009a)1 present data arguing against what they call "mainstream conventionalist" and "minimal conventionalist" accounts of the strengthening of scalar terms like *some*. Both positions (see Chierchia, Fox, & Spector, 2008 for a survey; see Geurts & Pouscoulous, 2009a, for additional references) claim that an "exclusivity" or **O**-operator is freely prefixed to any S node with the result that a proposition containing *some X, X or Y*, etc. is strengthened to 'some but not all,' exclusive or, etc.. Mainstream conventionalism claims that the strengthened interpretation is the preferred interpretation, unless it occurs in a context (eg. A downward-entailing context) which results in a logically weaker global interpretation of the sentence in which it occurs. Minimal conventionalism merely claims that the strengthened interpretation is possible, but says nothing about preference.

One way to evaluate conventionalist approaches is to examine 'embedded implicatures' (or, following Geurts & Pouscoulous, 'local scalar implicatures'). Consider a sentence like (1) (Geurts & Pouscoulous's (7a)):

**(1)**     All students read some of Chierchia's papers.

Insertion of the exclusivity operator under the scope of *all students* entails that all students read some but not all of Chierchia's papers and thus that no students read all of Chierchia's papers. This should be the preferred reading according to mainstream conventionalism, because it is a stronger (more limited) claim than the non-strengthened claim. It is also a possible reading according to minimal conventionalism. However, it is not a pragmatically justified reading from a Gricean perspective. The author of the statement presumably did not believe that all students read all of Chierchia's papers (else he would have said that). Thus,

---

Please address correspondence to: Charles Clifton, Department of Psychology, University of Massachusetts, Amherst, MA 01003, cec@psych.umass.edu, 413 545 2653.

[1]See Chemla, 2009, and Geurts and Pouscoulous, 2009b, for more discussion.

the pragmatically justified implication of (1) is (2a). It is not (2b), which is entailed if the exclusivity operator is inserted.

> **(2)**    a.    It is not the case that all students read all of Chierchia's papers.

> b.    All students read not all of Chierchia's papers.

Geurts and Pouscoulous (2009a) argue that introspective evidence is not adequate to decide what people usually do take sentences with scalar terms to mean (an argument that is particularly persuasive when the theorist is doing the introspecting). They present some very interesting 'verification' experiments which they claim disconfirm both flavors of conventionalism (but are consistent with a construal of Gricean pragmatics). In these experiments, a subject is shown a picture and asked whether a sentence containing a scalar term 'correctly describes' the picture. Their subjects nearly universally accepted sentences as correctly describing pictures that a strengthened interpretation of the sentence was not true of. For instance, 100% of Geurts and Pouscoulous's subjects accepted the sentence in Figure 1 (from Geurts & Pouscoulous 2009a) as correctly describing the arrangement shown in the figure, even though the locally-strengthened interpretation ('all of the squares are connected to some but not all of the circles' and thus 'none of the squares are connected to all of the circles') is false of the figure. They concluded, on the basis of data like these that "the conventionalist approach to scalar implicatures has little to recommend it" (Geurts & Pouscoulous, 2009a, p 431).

Geurts and Pouscoulous (2009a) acknowledged that data they obtained in verbal "inference" tasks (in which subjects are asked whether a sentence like *All the squares are connected with some of the circles* implies *All the squares are connected with some but not all of the circles*) exhibited a fair proportion (on the order of 50%) of strengthened interpretations. However, they state that such data are suspect. They argue that the proportion of acceptances of strengthened interpretations is inflated, perhaps because subjects' attention is called to the putative implication, so that subjects confuse it with the legitimate non-embedded Gricean implicature (*The square is connected with some of the circles* pragmatically implicates *The square is connected with some but not all of the circles*).

We were concerned that the verification task used by Geurts and Pouscoulous (2009a) has its own bias. Displays like that in Figure 1 can be correctly described in many ways: *There are squares and circles; Squares and circles are connected to each other; Some squares are connected to some circles*; etc. A pragmatic perspective does not require that only the strongest interpretation is a correct description, even if it is the preferred description. Similarly, while a mainstream conventionalist perspective claims that the preferred (strengthened) interpretation is not strictly true of the display, the existence of various weaker but legitimate descriptions of the display suggests that the non-strengthened interpretation may be acceptable. It may be that locally-strengthened interpretation is considered to be the best interpretation of the sentence, as long as it is the globally-strongest interpretation However, Geurts and Pouscoulous's subjects were not asked whether the display was the best possible depiction of the target sentence. They were only asked whether the sentence correctly described the display. A variety of weaker statements and interpretations can still be considered to be correct descriptions of the display.

From this perspective, it is tempting to consider what would happen if the subject were given a choice between two displays, one of which honors the locally-strengthened interpretation and the other of which violates it. If the locally-strengthened interpretation is the preferred one (as claimed by the mainstream conventionalist position), subjects should choose the display that honors it rather than the one that does not. If minimal conventionalism is on the right path, then subjects should be equally happy choosing either display. And the same should be true if Gricean pragmatics rules the day: the proper

interpretation should be 'All the squares are connected to some and possibly all of the circles.'

We conducted two experiments, modeled on Geurts and Pouscoulous's (2009a) Experiments 2 and 3. In each case, we shifted from a verification format to a choice format. Subjects were shown a sentence and two figures (generally one honoring a locally-strengthened interpretation, one honoring only a basic interpretation; see below for details), and asked to choose which picture was best described by the sentence: the 'strengthened' picture, the 'basic' picture, "both," and "neither." Both experiments were conducted in a single session, with randomly intermixed presentation of items including filler items, as described below.

## 1. Experiment 1

The first experiment was based on Geurts and Pouscoulous's Experiment 2, in which subjects were given a (Dutch) sentence like *Some of the B's are in the box on the left* and a picture containing the letters A, B, and C, and asked "to decide whether [the sentence] correctly describes [the picture]" (page 16). The left box had all the B's and all the A's, and the right box had all the C's. Geurts and Pouscoulous present this experiment not as a test of whether embedded implicatures are made (the sentences evaluated are simple sentences, presumably supporting the Gricean implicature that 'not all of the B's are in the box on the left') but simply as a check on the verification technique. They assumed that a subject who made the strengthened interpretation of *Some of the B's are in the box on the left* would reject that sentence as being a correct description of a picture where all the B's are in the box on the left. In addition to having their subjects verify whether such sentences correctly described the pictures, they had their subjects perform a written inference task. Subjects were asked to decide whether a sentence like *Some of the B's are in the box on the left* implies that not all the B's are in that box. 62% of their subjects accepted the truth of such a strengthened inference. However, a substantially smaller 34% of their subjects denied that the sentence correctly described the picture, as they should have done had they insisted on the strengthened interpretation.

The only claim that Geurts and Pouscoulous made for these data is that the inference technique yields inflated rates of scalar implicatures. We conducted Experiment 1 to shed light on whether this is the right claim, or whether the picture verification technique used by Geurts and Pouscoulous underestimated the incidence of scalar implicatures.

### 1.1. Materials

Four *some* sentences were constructed, as illustrated in (3). One pair of pictures was made for each sentence, as illustrated in Figure 2.

> **(3)**     Some of the stars are in the box on the left.

An additional 84 items (6 practice items plus 78 items from other experiments, including Experiment 2, presented below) were constructed. These were a mixture of picture verification items and written inference acceptance items, and tested both the scalar term *some* and the term *or*. We present only the *some* verification data here, for comparability with Geurts and Pouscoulous.

### 1.2. Subjects and Procedures

Thirty-six undergraduates at the University of Massachusetts participated; they received extra credit in their psychology courses in exchange for their participation. All subjects were tested individually. They viewed all the items on a computer monitor, and made their responses on a computer keyboard. The general instructions for all experiments were as follows:

In this experiment, you will be shown several short sentences. Following each sentence, there will be a question about the meaning of the sentence. On some trials, you will also be shown simple diagrams along with the sentences, and you will be asked to choose the diagram that is best described by the sentence. Please read the sentences carefully and answer each question to the best of your ability.

Subjects then advanced through 6 practice trials containing 3 simple verification and 3 inference items, followed by the individually-randomized presentation of a total of 82 experimental trials, including the 4 critical trials for Experiment 1. The verification instructions for all trials in all experiments simply asked subjects to 'Please indicate which shape is best described by the sentence below.' The sentence to be evaluated was presented below the verification instruction, and below the sentence was the diagram. The response options 'A', 'B', 'C (Both)' and 'D (Neither)' were indicated below the diagram (see Figure 2). Subjects made the verification response via key-press. No time constraint was imposed on the subjects, and participation in the study took approximately 20 minutes.

### 1.3 Results and Discussion

Table 1 contains the percentages of choices of each of the four options. The results are very clear. There was a preponderance of choices of the 'B' pair of boxes, in which some but not all of the named items (e.g., stars) were on the left; there were more choices of B than A: $t(35)=9.8$, $p<.001$, 95% CI of difference: (.54, .82). This, of course, is the choice that is consistent with a strengthened interpretation. Choices of 'both,' consistent with a non-strengthened 'some and possibly all,' were fairly infrequent and failed to rise above the arguable chance level of .25 choices of a given option, $t(35)=.13$, $p=.90$, 95% CI: (.13, .35). Choices of the A picture (which Geurts & Pouscoulous's subjects accepted 66% of the time) and the 'neither' item were essentially non-existent.

The methodological implication is clear: The verification task as used by Geurts and Pouscoulous (2009a) gives a much smaller estimate of the extent to which readers arrive at a strengthened interpretation of *some* in a non-embedded context than does the choice task we used. Geurts and Pouscoulous apparently assume that subjects will reject a sentence as a correct description of a picture if the most preferred interpretation of the sentence is not true of the picture. However, alternative interpretations of a sentence are possible; it is possible to cancel a scalar implicature. Under such an interpretation, the quantified sentence seems to be a possible description of the picture, permitting Geurts and Pouscoulous's subjects to accept it as such. However, our choice task permitted our subjects to let us know what their preferred interpretation of the quantified sentences is. They apparently took this opportunity to tell us, contrary to Geurts and Pouscoulous's conclusions, that they preferred the strengthened interpretation. This methodological conclusion justifies re-examining Geurts and Pouscoulous's verification results about the (non-) strengthening of embedded implicatures.

## 2. Experiment 2

The second experiment examined strengthening in embedded implicatures, using a task like that in Experiment 1. The critical items gave subjects a quantified sentence containing the scalar term *some* and asked them to indicate which of two displays it more accurately described, where one display pictured the 'some but not all' interpretation and the other pictured the 'all' possibility (see Figure 3, version 1; version 2 is a second type of test, described below). The basic predictions are as follows: If mainstream conventionalism is correct in a very strict sense, only the display that honors the strengthened ('some but not all') interpretations should be chosen. If minimal conventionalism is strictly correct, the "both" option should be chosen (and to the extent that a specific display is chosen, each

should be chosen equally often). The interpretation of a sentence strengthened by a conventional implicature is the denial of "all…all" (e.g., for the sentence *All the squares are connected to some of the circles*, it is 'It is not the case that all the squares are connected to all of the circles'). Since this interpretation is true of both the displays in the version 1 portion of Experiment 2, the pragmatic perspective predicts the same pattern of choices as minimal conventionalism does.

### 2.1. Materials

Four sentences were constructed that contained the scalar *some*. They were written in two versions each, as illustrated in (4), one with the universal quantifier *all* and the other with *each*.[2] Both forms involve embedded implicatures, and do not support scalar implicatures from a Gricean perspective. Each of the four items referred to a different triple of shapes.

**(4)** a. All of the squares are connected to some of the circles.

   b. Each of the squares is connected to some of the circles.

Two different figures, each with two designs, were made up for each of the four items. An illustration appears in Figure 3. One figure (top panel in Figure 3, Version 1) contained one design that honored the strengthened interpretation (the B item) and one design that honored the unstrengthened 'all' interpretation. The predictions for these items were laid out earlier. The other figure (bottom panel, Version 2) was designed so that neither design was true of the strengthened interpretation. For these items, a reader who arrived at that interpretation (i.e., a reader who made a local or embedded implicature) should choose Option D, 'neither.' A reader who did not take the strengthened interpretation should find either display acceptable and ideally choose Option C, 'both.'

### 2.2. Subjects and Procedures

Since they were conducted together, details regarding the subjects and procedures for Experiment 2 are identical to those of Experiment 1, with the exception that each subject received 8 critical trials. Each subject saw all four sentences twice, once where one figure honored the strengthened interpretation (Figure 3, Version 1) and once where neither figure did (Figure 3, Version 2). Two of each of these had the quantifier *all* and two, *each*, counterbalanced over subjects so that each item was tested with each quantifier equally often. Apart from this variation, trials differed only in the particular forms used (circles, triangles, stars, moons, hearts, etc.)

### 2.3. Results and Discussion

Table 2 contains the percentages of choices of each option. Trials on which subjects were presented with a design that honored the strengthened interpretation ('Version 1') provided evidence that they frequently arrived at the strengthened interpretation: There were substantial numbers of choices of the design that honored that interpretation, but essentially none of just the design that was inconsistent with it. t tests comparing the probability of a strengthened response to .25 indicated significant strengthening for Version 1, *t(71)=2.59, p<.05, 95% CI: (.28, .48)*. However, the most frequent choice was option 'C,' "both," which is the answer that is consistent with the non-strengthened, 'logical,' interpretation. Indeed, this option was chosen significantly more often than option B, *t(71)=2.13, p<.05, 95% CI of difference: (.01, .42)*.

---

[2]This manipulation was included based on the intuition – which proved to be incorrect – that the more individuating nature of *each* compared to *all* would discourage a 'group' interpretation of the predicate and encourage strengthening.

Trials on which neither design honored the strengthened interpretation received a substantially increased number of option 'D' ("neither") interpretations, which are consistent with a strengthened interpretation of the scalar, *t(71)=4.39, p<.001, 95% CI* of the difference: *(.10, .26)*. Version 2 also produced substantially more choices of option 'A' than option 'B,' *t(71)=4.15, p<.001, 95% CI of the difference (.11, .31)*, which is further reflected in a significant increase in the probability of choosing the A figure from Version 1 to Version 2, *t(71)=5.13, p<.001*, 95% CI of the difference: *(.15, .34)*. However, the most-frequent choice was option 'C,' "both," the interpretation that is consistent with the non-strengthened interpretation (vs. option A: *t(71)=2.74, p<.01, 95% CI: (.07, .43)*.

The greater frequency of choices of 'A' than of 'B' is of some interest. It has two apparent possible interpretations. From a Gricean perspective, a writer who wanted to describe the B picture would have written *Each of the squares is connected to all of the circles*. Since this is not what the sentence said, the sentence should not be taken to refer to the B picture. From a local strengthening perspective, the (strengthened) interpretation 'Each of the squares is connected to some but not all of the circles' is falsified by each of the squares in the B picture, but only by one square in the A picture. This could have encouraged choice of A as the 'less-wrong' alternative.

## 3. Conclusions

Methodologically, the conclusion is clear: While Geurts and Pouscoulous (2009a) may be correct in their concern that an inference judgment test yields an inflated number of instances of apparent strengthening of scalar terms, their alternative – the picture verification task, as they used it – apparently underestimates strengthening. When subjects were given a choice between two figures, only one of which honored the strengthened interpretation, they showed a distinct preference for choosing that figure. Geurts and Pouscoulous (2009a) took their verification data to show that subjects never, or almost never, rejected figures that violated strengthening of an embedded scalar term. Our data show that our subjects nonetheless showed a substantial preference for a figure that honored strengthening when given a choice between the two types of figures (and further, that they showed a smaller but still substantial frequency of rejecting both figures when neither honored strengthening) We submit that Geurts and Pouscoulous's conclusion that readers do not make embedded implicatures is based on suspect data, and hence is at best premature.

Theoretically, though, the cup may be only half full. While our data show that readers who make the choice between the strengthened and the unstrengthened interpretation of an embedded scalar strongly prefer the former, they also show that the most common response is not to choose between the interpretations but to accept both. Such ecumenism is not a given; Experiment 1, which tested non-embedded scalar terms, found that "both" choices were fairly infrequent. The choice of "both" in Experiment 2 presumably reflects the absence of strengthening. Perhaps the right conclusion is that an apparently strengthened interpretation of an embedded scalar term like *some* is possible, but not obligatory and not even preferred. This conclusion may present some difficulty to one who holds a pragmatic Gricean perspective. As Geurts and Pouscoulous (2009a) make clear, Gricean accounts of strengthening of scalar terms under the scope of (e.g.) *think* and *believe* (Guerts, 2009) do not readily generalize to scalar terms under the scope of *all* or *each*. In the absence of a Gricean account of pragmatic strengthening under the scope of such terms, our results call Gricean accounts generally into question. Similarly, our findings may present some difficulty for a mainstream conventionalist perspective: It is not clear from such a perspective why the strengthened interpretation is apparently taken less frequently than the basic interpretation. The minimal conventionalist perspective discussed by Geurts and Pouscoulous (2009a) can accommodate our data, as can a perspective that says that terms

like *some* are simply ambiguous, but these perspectives are so unconstraining that one would hope to adopt them only as a last resort. We can conclude only that the evidence presented by Geurts and Pouscoulous (2009a) has not made a solid case against the existence of local, embedded implicatures. We trust that additional experimental research will clarify the conditions under which such implicatures are made, and hope that additional linguistic analysis will shed light on why these conditions encourage strengthening.

## References

Chemla E. Universal implicatures and free choice effects: Experimental data. Semantics and Pragmatics. 2009; 2(Article 2):1–33.

Chierchia, G.; Fox, D.; Spector, B. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In: Maienborn, C.; Heusinger, Kv; Portner, P., editors. Semantics: An International Handbook of Natural Language Meaning. Berlin: Mouton de Gruyter; 2008. (in preparation; URL http://semanticsarchive.net/Archive/WMzY2ZmY/CFS_EmbeddedSIs.pdf

Geurts B. Scalar implicature and local pragmatics. Mind and Language. 2009; 24:51–79.

Geurts B, Pouscoulous N. Embedded implicatures?!? Semantics and Pragmatics. 2009a; 2(Article 4): 1–34.

Geurts B, Pouscoulous N. Free choice for all: A response to Emmanuel Chemla. Semantics and Pragmatics. 2009b; 2(Article 5):1–10.
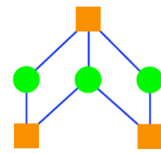
All the squares are connected
with some of the circles.

□ true    □ false

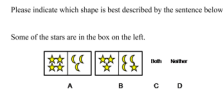**Figure 1.**
(From Geurts and Pouscoulous)

Please indicate which shape is best described by the sentence below

Some of the stars are in the box on the left.



**Figure 2.**
Illustrations of figures used in Experiment 1

Version 1. Figure used where B option illustrated the strengthened interpretation

Please indicate which shape is best described by the sentence below

All/Each of the squares are connected to some of the circles.

Both    Neither

A        B        C        D

Version 2. Figure used where neither option illustrated the strengthened interpretation

Please indicate which shape is best described by the sentence below

All/Each of the squares are connected to some of the circles.

Both    Neither

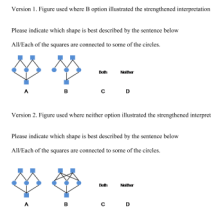A        B        C        D

**Figure 3.**
Illustrations of figures used in Experiment 2

**Table 1**

Percentages of choices of each option (standard errors in parentheses), Experiment 1.

| | Choice Option | | |
|---|---|---|---|
| A | B[*] | C ("both") | D ("neither") |
| 3 (2) | 71 (6) | 24 (5) | 2 (2) |

"Strengthened" alternative indicated by *

## Table 2

Percentage of choices of each option (standard errors in parentheses).

| Version 1: B alternative strengthened | | | | |
|---|---|---|---|---|
| **Quantifier** | **A** | **B** | **C** | **D** |
| Items with one design that honors strengthened interpretation | | | | |
| *all* | 3 (2) | 39* (7) | 57 (8) | 1 (1) |
| *each* | 0 (0) | 38* (7) | 63 (7) | 0 (0) |
| Version 2: Neither alternative strengthened | | | | |
| Items with no design that honors strengthened interpretation | | | | |
| *all* | 28 (7) | 6 (3) | 50 (8) | 17* (6) |
| *each* | 24 (6) | 4 (2) | 51 (8) | 21* (6) |

Strengthened choice indicated by *.