Scalar Implicature and Local Pragmatics

BART GEURTS

Abstract: The Gricean theory of conversational implicature has always been plagued by data suggesting that what would seem to be conversational inferences may occur within the scope of operators like *believe*, for example; which for bona fide implicatures should be an impossibility. Concentrating my attention on scalar implicatures, I argue that, for the most part, such observations can be accounted for within a Gricean framework, and without resorting to local pragmatic inferences of any kind. However, there remains a small class of marked cases that cannot be treated as conversational implicatures, and they do require a local mode of pragmatic interpretation.

1. Introduction

Even before Grice's Harvard Lectures had started appearing in print, his critics were arguing against his theory of conversational implicature using examples in which, prima facie, conversational implicatures occurred within the scope of conditionals and other operators; which on Grice's view amounts to a contradiction in terms, since conversational implicatures can only be derived on the basis of a full-blown speech act (Cohen, 1971). For many years, this discussion dragged on without too much fervour, but then it became livelier when Landman (1998), Levinson (2000), and Chierchia (2004) began calling for drastic departures from the Gricean party line, provoking revisionist responses from Sauerland (2004), van Rooij and Schulz (2004), Spector (2006), and Russell (2006), among others. The current stage of debate concentrates, as does this paper, on the status of one particular type of Gricean inference known as 'scalar implicature'.[1]

Scalar implicatures are supposed to work as follows. Clyde says:

    (1)   Bonnie had some of the pears.

Address for correspondence: Department of Philosophy, University of Nijmegen, Postbox 9103, 6500 HD Nijmegen, The Netherlands.
Email: bart.geurts@phil.ru.nl

[1]   Terminological note: my use of the word 'Gricean' is somewhat of a compromise. Since Grice never had much to say about scalar implicatures, I would have preferred 'neo-Gricean', but this term has become effectively useless, as it is commonly applied to radically different doctrines. (In particular, Horn and Levinson are neo-Griceans, but only the former is a Gricean, at least as I understand the term.) Moreover, for the purposes of this paper, 'Gricean' is not in opposition to 'relevance theoretic' (Sperber and Wilson, 1995). In itself, relevance theory is consistent with what I call the Gricean account of scalar implicature, and the only sustained relevance-theoretic discussion of scalar inferences that I'm aware of accepts it as part of the story (Noveck and Sperber, 2007).

On the classical Gricean account, (1) means that Bonnie had at least some of the pears, and may implicate that she didn't have all of them. (Some authors—not I—would say that this implicature is standardly associated with the sentence.) This implicature is explained by assuming that the hearer reasons, and is entitled to reason, as follows:

   i.   Rather than saying (1), Clyde could have said:

        (1*) Bonnie had all the pears.

        Why didn't he do so?
   ii.   The most likely explanation is that Clyde doesn't believe that (1*) is true: $\neg B_c(1^*)$.
   iii. Clyde is likely to have an opinion as to whether (1*) is true: $B_c(1^*) \vee B_c\neg(1^*)$.
   iv.   Between them, (ii) and (iii) entail $B_c\neg(1^*)$: Clyde believes that Bonnie didn't have all the pears.

Actually, this derivation involves not one but two implicatures, one of which $(\neg B_c(1^*))$ is weaker than the other $(B_c\neg(1^*))$.[2] In order to proceed from the former to the latter, we use the premiss that the speaker is not undecided about the truth of the stronger alternative. In most cases, this is just to say that the speaker knows the relevant facts; which is why, following van Rooij and Schulz (2004), I call it the 'competence assumption'.[3] If the competence assumption doesn't hold, the strong implicature cannot be derived, either.

   Although it is tempting to view this kind of analysis as a set procedure for generating implicatures, doing so would go against the true Gricean spirit. The derivation of a scalar implicature starts with the question why the speaker didn't make a stronger statement than he actually did, and the kind of answer given above presupposes that the speaker's reasons are epistemic in nature, that they have to do with what he knows or believes. I assume that this presupposition is justified often enough to treat it as the normal case. But it isn't always so. For example, it

---

[2]   Here I follow the standard line in assuming that ¬Bc(1*) is the right way of rendering the first implicature. Pouscoulous (2006, pp. 43–46) argues against this, claiming that the first implicature is, rather, ¬Bc(1*) ∨ ¬Bc¬(1*), i.e. for all Clyde knows, Bonnie may or may not have had all the pears. I agree with her, and consequently I believe that, as presented here, the Gricean analysis is strictly speaking false. But since Pouscoulous' amendment is orthogonal to my present purposes, I hope I will be excused for adopting the conventional account, warts and all, because it helps to simplify the discussion somewhat. In the same vein, as nothing will hinge on the distinction between speakers' knowledge and belief, I will not try to keep these notions separate, and in fact will be using whichever is the least misleading at any given point.
[3]   As pointed out to me by Larry Horn, this style of disjunctive reasoning was already used by Bartsch (1973) to explain 'neg-raising' in propositional-attitude verbs. See Horn (1989) for further discussion.

might be common knowledge that Bonnie is deeply ashamed of her fruit addiction, and would be furious if she learned that Clyde had said (1*), even if it was true. In such a situation, Clyde's uttering (1) might convey that he didn't feel at liberty to make a stronger statement. This would be an implicature, too, but it seems less typical, and I will leave it aside for most of this paper.

The problem of local implicatures is illustrated by the following case, in which Clyde says:

(2)   Prentiss believes that Bonnie had some of the pears.

Intuitively, (2) may well give rise to the implicature that, according to Clyde, Prentiss believes that Bonnie didn't have all of the pears. Prima facie, it seems as if in this case the scalar implicature is derived within the scope of an attitude verb. This causes an embarrassment to the Gricean account, as becomes clear when we deploy the same argument that worked so well for (1):

i.   Rather than saying (2), Clyde could have said:

(2*)   Prentiss believes that Bonnie had all the pears (or $B_p(1^*)$ for short).

Why didn't he do so?

ii.   The most likely explanation is that Clyde doesn't believe that (2*) is true: $\neg B_c B_p(1^*)$.

iii.   Clyde is likely to have an opinion as to whether (2*) is true: $B_c B_p(1^*) \lor B_c \neg B_p(1^*)$.

iv.   Between them, (ii) and (iii) entail $B_c \neg B_p(1^*)$: Clyde believes it isn't the case that Prentiss believes that Bonnie had all the pears.

But this isn't good enough: the predicted implicature ($B_c \neg B_p(1^*)$), though not incorrect, is weaker than what we would like to have ($B_c B_p \neg (1^*)$).

Cases like (2) have led some authors to argue that scalar implicatures aren't really implicatures. Rather, according to them, upper-bounding inferences are more or less directly associated with scalar expressions, like *some*, *or*, *warm*, and so on. Not to put too fine a point on it, as yet, the idea is that *some* actually *means* 'some but not all'. If this is right, the 'not all' inference associated with *some* goes directly into the Fregean content of the sentence at the point where the word occurs; and if the word is sitting in the scope of *believe*, say, that is where the 'not all' inference will be triggered, too. For obvious reasons, I will call this type of view 'localist'.

Despite the fact that some localists profess to be neo-Griceans (Levinson is a prominent case in point), it will be evident that localism entails a radical departure from Gricean principles. Since it is generally agreed, even by localists, that these principles (or something similar) are needed anyway, we should ask ourselves whether such a departure is really necessary. That is what I will do in the remainder of this

paper, which is organised as follows. To begin with, I present the evidence that has been used to argue for localism and against the Gricean approach to scalars (§2). I separate the recalcitrant data into two groups: L–type and C–type. Initially, this distinction is motivated solely by the fact that Levinson is mainly concerned with the first class, while Chierchia confines his attention to the second; but then I proceed to argue that the division goes deeper than that. In §3, I briefly discuss the various localist ways in which these problematic data might be accounted for. With these preliminaries out of the way, the dialectical part of the paper begins to unfold, in which I argue first that, if conceived as a general theory of scalar inference, localism runs into all sorts of trouble (§§4-5). Then I try to show that, first impressions notwithstanding, the Gricean framework deals quite well with run-of-the-mill cases of local implicature, i.e. the C–type cases (§6), though not with the marked L–type cases (§7). The inexorable conclusion of my argument is threefold:

- There are two quite different kinds of upper-bounding scalar inferences: marked (L–type) and unmarked (C–type).
- The Gricean approach is basically correct, in the sense that it accounts for all the unmarked cases, and does it better than any localist theory could ever do.
- The marked cases have nothing to do with conversational implicature, and it is these cases, and these cases alone, that involve a genuinely localist mode of interpretation.

Recently, researchers on the banks of the Charles River have begun to advertise a wholesale syntactisation of scalar implicature (Chierchia, 2006; Fox, 2006). For most of this paper I will ignore this unfortunate development, but I will briefly discuss it towards the end (§8), and argue that it isn't just wrong but pointless, too: since scalar inferences can be explained entirely in terms of independently motivated pragmatic principles, any attempt at syntactic explanation is bound to be futile.

2. Two Kinds of Scalar Inference

In this section I present the observations that have been used to argue against the Gricean approach and for localist analyses of all stripes. The question whether or not they have been rightly used for these purposes will be addressed in §6; until then, the data will be taken at face value. I will sort the examples into two groups, labelled 'L–type' and 'C–type' to reflect that they have been used by Levinson and Chierchia as key evidence for their respective brands of localism. For the time being, I will not attempt to justify the division, but later on I will argue that it actually carves nature at its joints. Starting with the L–type data, let us have a look at comparatives first. The following examples are Levinson's (2000, pp. 203, 204):

(3) a.   Drinking warm coffee is better than drinking hot coffee.
    b.   A teacher who is sometimes late is preferable to one who is always late.

If the lexical meaning of *warm* is 'at least warm', then the Fregean meaning of (3a) is that drinking warm-and-possibly-hot coffee is better than drinking hot coffee. Pre-theoretically speaking, this makes little sense, and theoretically speaking, too, it isn't likely to be right. It seems reasonable to assume that comparisons are always between non-overlapping possibilities, and with the hypothesised meaning of warm they aren't. It is hard to see, therefore, how we can avoid the conclusion that, in this particular case at least, *warm* means 'warm but not hot'. The same, *mutatis mutandis*, for *sometimes* in (3b). Another type of example used by Levinson to buttress his localist view involves conditionals like the following (these are due to Horn (2006, p. 27) who, though not a localist himself, has nicer examples than Levinson):

(4) a. If it's warm, we'll lie out in the sun. But if it's VERY warm, we'll go inside and sit in front of the air-conditioner.
     b. If you're convicted of a felony, you'll spend at least a year in jail. And if you're convicted of murder, you'll be executed.

On the face of it, if the meanings of *warm* and *very warm* weren't disjoint, (4a) should fail to make sense; but in fact the discourse, if somewhat marked, is perfectly felicitous. Hence, we are forced to conclude that, in the context of (4a), *warm* means 'warm but not very warm'; and similarly, in the context of (4b), the meaning of *felony* should exclude murder.

The third and last variety of L-type examples (from Horn 1989: 382) involve negation:

(5) a. Around here, we don't LIKE coffee, we LOVE it.
     b. I'm not HAPPY he's gone—I'm ELATED.

Suppose that negation in English is unambiguous and that it means the same thing as in classical logic. Then it is hard to see how we can avoid concluding that, in (5a), the Fregean meanings of *like* and *love* exclude each other, and that the same goes for *happy* and *elated*, as these words are used in (5b).[4]

Turning to the C-type cases, let us first consider attitude verbs like *believe* and *know* (examples by Chierchia, 2004, pp. 44, 45):

(6) a. John believes that some students are waiting for him.
     b. John knows that some students are waiting for him.

---

[4] Alternatively, it might be suggested that the examples in (5) are instances of 'metalinguistic negation' (Horn, 1985, 1989). For reasons discussed at length (and ad nauseam) by McCawley (1991), Carston (1996), and myself (1998a), I reject this option. I will briefly return to this issue in §7.

As discussed in the introduction, scalars occurring in the scope of *believe* seem to give rise to local upper-bounding inferences: (6a) may be read as 'John believes that some but not all students are waiting for him.' Factive *know* is somewhat different. The critical reading of (6b) is the one on which the sentence implies that not all of John's students are waiting for him. Chierchia explains this by assuming that the reinforced meaning of the complement clause goes into the presupposition triggered by *know*; hence the sentence presupposes that some but not all of John's students are waiting for him. Crucially, this explanation hinges on the assumption that the upper-bounding inference is local.

The examples in (6) are problematic for the Gricean theory because it yields predictions that are too weak. The following examples are potentially more serious, for in these cases the Gricean approach seems to make incorrect predictions:

(7)   a.   Mary is either working at her paper or seeing some of her students (Chierchia, 2004, p. 46).
       b.   Kai had the broccoli or some of the peas last night (Sauerland, 2004, p. 370).

Let's consider (7b) ((7a) is similar). One of the stronger alternatives available to someone who utters (7b) is (7b*):

(7)   b.* Kai had the broccoli or all of the peas last night.

If the Gricean account is right, we should expect that (7b) can give rise to the implicature that, according to the speaker, (7b*) is false, and therefore that Kai didn't have the broccoli. But this prediction is patently wrong.

A second type of case in which Gricean reasoning seems to yield incorrect results is the following (Chierchia, 2004, p. 46):

(8)   A: How did students satisfy the course requirement?
       B: Some made a presentation or wrote a paper. Some took the final test.

One of the stronger alternatives to B's first utterance is:

(8)   B*:   Some made a presentation and wrote a paper.

If the Gricean account were right, we should expect that (8B) may give rise to the implicature that, according to the speaker, (8B*) is false, hence that there were no students who made a presentation and wrote a paper. But this prediction seems too strong. In both cases, (7) as well (8), localist predictions appear to be better than Gricean ones.

3. Varieties of Localism

There are various ways of implementing the basic localist idea that *warm* may be used to mean 'warm but not hot'. Some localists analyse scalars after the model of expressions like *six pizzas later* or *marble lion*. Literally speaking, these expressions are oxymoronic, but in context they are *re*interpreted in such a way that they make sense; that is to say, their Fregean content is edited so as to fit the context: *six pizzas later* comes to mean something like 'having eaten five pizzas', and in the context of *marble lion*, the word *lion* is reconstrued as 'lion statue'. The view that the same process may serve to give *warm et al.* their upper-bounded meanings has been defended by Bach (1994) and myself (1998b), among others.

Another way of looking at the 'at least some' and 'some but not all' interpretations of *some* is that the word's lexical content is indeterminate between the two readings (Bezuidenhout, 2002; cf. also Carston, 1998 on number words). So the conventional content associated with *some* is neither 'at least some' nor 'some but not all', but is neutral between the two. For reasons I have discussed elsewhere I doubt this will work (Geurts, 2006c), but that is as it may be. The main point, for current purposes, is that such views, too, are localist.

Chierchia (2004) proposes a two-dimensional semantic calculus, with ordinary Fregean meanings living in one dimension, and reinforced meanings in the other. Reinforced meanings are computed in much the same way as Fregean meanings. The main differences are that the reinforced meaning of a scalar expression contains an upper-bounding component (i.e. in this dimension *good* means 'good but not outstanding'), which is suppressed in downward-entailing environments like (9b):

(9) a. This champagne is good.
     b. This champagne is not good.

While the reinforced meaning of (9a) is that the champagne is good but not outstanding, in (9b) the upper-bounding component of *good* is suppressed, so that the reinforced meaning of the sentence comes out the same as its Fregean meaning: in either dimension, the latter sentence entails that the champagne is less than good, rather than either less than good or outstanding (which would be the negation of 'good but not outstanding').

Yet another way of fleshing out localism is by stipulating that the grammar provides a covert operator which basically acts as a silent *only* that optionally associates with a scalar expression so as to deliver an upper-bounded construal.[5] According to this variant of localism, (10a) may be interpreted, in effect, as (10b):

---

[5] Onlification theories are defended by Chierchia (2006) and Fox (2006), both of whom stipulate that silent *only* need not be interpreted in situ, thus adding greatly to the power of this type of account (see §8 for discussion). Note that when I speak of 'localist' theories, this option isn't included. Hence, while Chierchia's 2004 theory is strictly localist, its 2006 successor is not.

(10)   a.   Some of the goats are sick.
       b.   Only some of the goats are sick.

Although I sympathise with the first view, according to which (some) scalar inferences are to be analysed in terms of pragmatic meaning shifts, I will for the time being ignore the differences between these various ways of practising localism. For the purposes of this and the next few sections, the key feature common to all these ways is that they allow operators to see the upper-bounding inferences associated with scalar expressions occurring in their scope. This, and only this, is what I mean when I say that, on localist accounts, *some* is interpreted as 'some but not all'.

There is another distinction that I cannot afford to ignore, viz. between defaultist and noncist varieties of localism. If you are a defaultist, you are committed to the view that scalar expressions give rise to upper-bounding inferences as a matter of course; it is what will happen normally. This description is admittedly vague, but then the doctrine itself isn't always easy to nail down, as witness the following passage:

> By default interpretation, I simply mean the one that most people would give in circumstances in which the context is unbiased one way or the other (Chierchia, 2004, p. 51).

Since every context is 'unbiased one way or the other', I'm not at all sure what this says, but the author's intended meaning is clear enough: *ceteris paribus*, scalar expressions give rise to upper-bounding inferences. Chierchia and Levinson are both defaultists in this sense of the word.[6] Other authors adopt the position that scalar inferences are entirely dependent on the context; they are noncists.

Note that I'm using the term 'default' in a weaker, and therefore more general, sense than it has been used in the recent experimental literature, which has successfully lambasted the stronger view according to which scalar implicatures are automatic inferences (e.g. Bott and Noveck, 2004; Breheny *et al.*, 2006). Levinson endorses the stronger view, but Chierchia is not committed to it.

From an aesthetical point of view, it is somewhat unfortunate that the neat divisions laid out in the foregoing are blurred by the existence of hybrid theories. In fact, these theories are so common that hybridism is practically the norm. Thus, Recanati (2003) proposes to incorporate Chierchia's default theory into a localist model of the noncist sort, and while Bach (1994) and Noveck and Sperber (2007) defend the position that scalar inferences are usually nonce inferences, but on

---

[6]   The defaultism/noncism distinction is orthogonal to the opposition between localist and Gricean theories. Horn is a Gricean of the defaultist persuasion; younger exponents of the Gricean school, like van Rooij and Schulz and the author of Spector, 2006, are noncists. As we will presently see, the choice between defaultism and noncism is more critical for localists than for Griceans.

occasion may be genuine implicatures, for Horn (2006) it is the other way round: upper-bounded construals are usually implicatures, but sometimes they may issue from local nonce inferences; and so on. In the end, I will be pleading for a hybrid theory, myself: the following discussion will culminate in the conclusion that Horn is right.[7] That is to say, Grice's is the best general framework for explaining scalar inferences, though a modicum of localism is needed for dealing with some out-of-the-way problem cases.

4. Defaultism

Methodologically speaking, noncist versions of localism are disappointingly weak in the sense that they merely predict that local inferences may or may not occur. Clearly, this type of theory needs to be complemented by an account of why and when scalar inferences arise. As far as I am aware, there is no such theory on the market. Hence, one reason for paying extra attention to defaultist varieties of localism is that, unlike their noncist siblings, they make substantial predictions. Another is that defaultism is an especially influential brand of localism.

Towards the end of this paper, I will argue that there is a place for localism in a general theory of scalar inference, but that its status is rather marginal: there is a truly local species of scalar inference, but it is a deviant phenomenon that only occurs in narrowly circumscribed circumstances. All localist theories that have been propounded in the literature are more ambitious than that, and consequently they have little predictive bite unless they are defaultist, as well. In other words, although localism and defaultism aren't wedded to each other, there is a natural affinity between the two, and the theories advocated by some of the most prominent localists, like Chierchia and Levinson, espouse defaultism, as well. In this section, I will argue, to begin with, that there are strict limits to what any defaultist can reasonably hope to explain, and then I will go on to argue that defaultism is a lost cause.

4.1 Defaultism and Markedness
It sounds truistic to say that default rules can't explain exceptions, but oddly enough Levinson seems to think otherwise. Consider the following example:

(11) a. It isn't likely that the match will be cancelled.
      b. It isn't likely that the match will be cancelled: it's certain.

---

[7] This is not to say that I agree with Horn across the board. For one thing, Horn is a defaultist, which I am not. For another, Horn is a better Gricean than I in that he sees his theory of implicature as part of an account of speaker's meaning rather than utterance interpretation, whereas for me it is chiefly the latter.

Everybody agrees that the standard, unmarked interpretation of (11a) is that there probably will be a match. However, as is well known, the same sentence can be used, as in (11b), to help convey that the match is cancelled. The latter interpretation is strongly marked: it typically requires an explicit contrast to enforce this construal.

In its crudest form, defaultist localism entails that the bread–and–butter interpretation of (11a) denies that it is likely but not certain that the match will be cancelled. But this is the reading we observe in (11b), which is the marked case. Therefore, a more sophisticated variety of defaultism is called for, which is provided by Chierchia (2004). He proposes a grammatical procedure designed to ensure that scalar inferences are suppressed in downwardentailing environments, and predicts that, in (11a) as well as (11b), the scalar inference associated with *likely* is cancelled. Thus, Chierchia gets the right predictions for (11a), though not for (11b).

Levinson's position is a puzzling one. He uses strongly *marked* examples like (11b) as evidence for his view that scalar inferences are triggered locally *and by default*, which makes no sense at all, as far as I can see. What is more, all the examples used by Levinson to prove that scalar inferences affect Fregean content turn out to be marked in much the same way as (11b) is. Let me quote just one of Levinson's own examples:

(12)   If the chair sometimes comes to department meetings that is not enough; he should come always (Levinson, 2000, p. 205).

Normally speaking, scalar inferences do not arise in the antecedent of a conditional, since this position is downward entailing. In this particular case, however, it seems that *sometimes* has to be interpreted as 'sometimes but not always'. Levinson is surely right about this. But he cannot have his defaultist cake and eat it. If you choose to be a defaultist, then your theory *should not* account for the readings observed in (11b) and (12).

4.2 The Case Against Defaultism

The myth that scalar expressions prompt upper-bounding inferences by default is probably an artefact of the method that is most commonly used for collecting linguistic data: introspection.[8] If we ask ourselves whether we would take an utterance of (13a) to imply that, according to the speaker, (13b) holds as well, it seems clear that we would:

---

[8]   I should note that I have no problems with the introspective method as such. I have been using it for quite a while, and have no intention of giving it up. My point is just that it is a biased tool when it comes to gathering data on conversational implicatures.

(13) a. Some of the goats have a cough.
      b. Not all of the goats have a cough.

However, this introspective method of collecting data on implicature is arguably biased (Geurts, 2006a; Geurts and Pouscoulous, 2008). Obviously, to ask oneself whether or not (13a) implies (13b) is to suggest already that it might be implied, but more importantly, this question raises the issue whether or not all of the goats have a cough, or in other words, it makes it relevant to establish whether this is the case. And even if our intuitions about this case are dependable (which is as it may be), they do not necessarily tell us anything about how (13a) is interpreted in situations where this issue is not at stake. In short, the introspective evidence used to buttress claims to the effect that scalar inferences are defaults should not be taken at face value.

In order to test our intuitions on the default status of scalar inferences, quantitative data are needed. In particular, we need experiments that test for scalar inferences in a context that is as neutral as possible. It turns out that the psychological literature, and especially the literature on human reasoning, contains a number of studies that meet this requirement; I will give one example here. In an acquisition study, Paris (1973) presented his participants with disjunctive sentences with arbitrary content, such as the following:

(14) The bird is in the nest or the shoe is on the foot.

Paris's materials contained sentences with *or* as well as *either… or*. Participants were asked to determine whether or not such sentences were true of a pair of pictures. Paris's main finding was that, overall, inclusive interpretations were preferred for 82% of the *or*-sentences and 76.5% of the *either…or*-sentences. (The difference between *or* and *either…or* was significant, though much smaller than one might have expected.) For the adult participants, the rates were 75% and 68.5%, respectively. These data suggest that the 'normal' interpretation of *or* is inclusive, and therefore go against the defaultist view.

Paris's results are typical of what one finds in the literature: once contextual factors are factored out and the experimental paradigm is as neutral as possible, rates of scalar inferences are typically around chance level, give or take 10%. Paris's results on *or* are confirmed, with a quite different experimental design, by Evans and Newstead (1980). For parallel results on *some*, see e.g. Begg and Harris (1982), Newstead and Griggs (1983), Bott and Noveck (2004), and Noveck (2001), who also presents data on *might* that show the same pattern. In short, the bulk of the relevant experimental evidence on scalar terms falsifies the claim that these expressions prompt upper-bounding inferences by default.

4.3 Summing Up
A localist theory needs an account of the circumstances under which scalar inferences will or will not occur. Defaultism offers such an account, but it has two

problems. First, it has intrinsic limitations: defaultism cannot explain marked, L-type instances of scalar inference. Secondly, and more seriously, experimental evidence suggests rather forcefully that defaultism is wrong. Hence, the only viable form of localism is a noncist one, and it remains an open question why scalar inferences do or do not arise in any given situation. In §7 I will suggest an answer to this question, and thus vindicate localism, after a fashion, but this will be a Pyrrhic victory, because if I am right the best localism can hope for is a marginal role in the theory of scalar inference.

5. More Problems with Localism

In the last section I argued that, if it aims at offering a general account of scalar inference, localism is faced with the awkward dilemma of being either ineffectual (noncism) or wrong (defaultism). This section takes a different tack. I will try to show that, no matter whether noncist or defaultist, any version of localism is bound to fail at explaining some phenomena (§§5.1-5.3), and is falsified by others (§5.4).

5.1 The Variability of Quantity Implicatures

As discussed in the introduction, the Gricean account distinguishes between several kinds of implicature that may be associated with any scalar expression. If I utter (15a), I may variously be understood as implying (15b), (15c), or (15d):

(15)   a.   Some of my cousins live in Belgium.
       b.   The speaker is not at liberty to say whether all his cousins live in Belgium.
       c.   The speaker doesn't believe that all his cousins live in Belgium.
       d.   The speaker believes that not all his cousins live in Belgium.

On the Gricean account, some of these possibilities hang together: if my utterance licenses the inference that (15d) is the case, it is because it licenses the weaker inference in (15c) first, which in conjunction with the competence assumption yields (15d). Localist theories only predict the last inference. They don't take the other possibilities into account, and a fortiori they don't connect the strong implicature (15d) with the weak one (15c); rather, their claim is that the strong implicature is always generated directly.[9] Hence, localism misses out on what appears to be a natural connection, i.e. between weak and strong implicatures, but

---

[9]   To be sure, it is not a metaphysical necessity that localism should confine its attention to (15d). Technically speaking, it would be possible to maintain that (15a) is multiply ambiguous, and that each of the 'readings' in (15b-d) springs from the lexical content of *some*. Nobody has ventured such a claim so far, and for good reason: not only would it involve unprecedented complications of the grammar, it would also be blatantly ad hoc.

more importantly, it fails to account for a range of inferences that, consequently, will have to be explained otherwise, and that look very much as if they can only be explained on Gricean principles.

5.2 Strong Non-local Implicatures

We have just seen that localism fails to account for some scalar inferences that are presumably different in nature from strong upper-bounding inferences of the 'some but not all' type. However, it also fails to account for some inferences that would seem to be of this type:

(16)   a.   You may have an apple or a pear $\rightsquigarrow$ You may not have both.
     b.   You may have some of the apples $\rightsquigarrow$ You may not have all of them.
     c.   She may have eaten some of the apples $\rightsquigarrow$ She will not have eaten all of them.

Each of these sentences is of the form $\diamondsuit[\ldots \alpha \ldots]$, and may give rise to the inference that $\neg\diamondsuit[\ldots \beta \ldots]$, where β is any of α's stronger scale mates. Whereas the Gricean theory explains these inferences without further ado, localist theories yield predictions that are too weak. Consider (16a), for example (the other cases are analogous). The Fregean content of this sentence is of the form $\diamondsuit[\varphi \vee \psi]$; its reinforced content, with a localist treatment of scalar *or*, would be $\diamondsuit[[\varphi \vee \psi] \wedge \neg[\varphi \wedge \psi]]$. But instead of entailing that the addressee isn't allowed to have an apple and a pear, this entails merely that he is allowed not to have an apple and a pear. Hence, the localist predictions about (16a–c) fall short of their explananda.

5.3 Multiplicatures

As argued in Geurts (2006b), some implicatures are derived not on the basis of a single proposition but from propositional ensembles. The following is a case in point:

(17)   Q: Which cities did you see on your trip to Italy?
     A: Clyde and I first went to Naples and Rome together. Then, while he went to see Venice, I visited Florence.

Bonnie's answer suggests rather strongly that, on this trip, she saw only Naples, Rome, and Florence. The Gricean theory accounts for this in the familiar way: The purpose of Bonnie's discourse is to inform the addressee about the places she visited on her Italian tour, and she lists Rome, Naples, and Florence. If Bonnie had been to Milan, say, she would be withholding relevant information, and assuming she is a cooperative speaker, she wouldn't do that. Therefore, she didn't visit Milan. It will be evident that there is no way a localist theory could account for inferences like this.

The Gricean approach allows us to treat multiplicatures in the same way as scalar implicatures. In both cases, the hearer reasons that the speaker doesn't believe that φ because if he had he would have said φ. That this idea is a feasible one is proved by recent formalisations of Gricean reasoning, which allow for a uniform treatment of scalar implicatures and multiplicatures (van Rooij and Schulz, 2004; Spector, 2006). Of course, it might still be doubted that this is right, and it is clear that proponents of a localist view will have to deny that it is right. But it is equally clear that the burden of proof is on them, and as far as I can tell there are no good reasons for believing that scalar implicatures are unrelated to other kinds of upper-bounding inference.

5.4 Non–implicatures

I have given a number of examples which show that localist theories fail to account for a range of facts that aren't particularly problematic for the standard Gricean theory. But even if it doesn't go far enough, it might still be possible that the localist approach is correct as far as it goes. Cases like the following rule out that possibility, too:

(18) a. I prefer to visit Tokyo or Kyoto $\not\leadsto$ I prefer not visiting both.
  b. I hope that some of my relatives will remember my birthday $\not\leadsto$ I hope that not all of them will remember it.
  c. Bonnie said that Clyde bought a new car or bicycle $\not\leadsto$ She said he didn't buy both.

(19) a. All the farmers in this region own goats or sheep $\not\leadsto$ None of them own both.
  b. At least 300 of the farmers in this region own goats or sheep $\not\leadsto$ At least 300 of them don't own both.
  c. There are exactly 300 farmers in this region who own goats or sheep $\not\leadsto$ There are exactly 300 farmers in this region who don't own both.

Contrary to what localist accounts predict, it is clear that, e.g., (18a) would not normally imply that the speaker would prefer not visiting Tokyo and Kyoto. In fact, good specimens of unmarked local upper-bounding construals appear to be exceedingly rare. This impression is confirmed by a study I conducted with Nausicaa Pouscoulous (Geurts and Pouscoulous, 2008), in which we asked subjects to assess inferences like the following:

(20) a. Bonnie has to eat some of the apples $\leadsto$ She isn't allowed to eat all of them.
  b. Every boy had an apple or a pear $\leadsto$ None of them had both.

The rates at which our subjects endorsed these inferences were 3% and 27%, respectively. If we take into consideration that this experimental paradigm is arguably biased in favour of finding implicatures (as I argued in §4.2), these results entail that localist inferences are virtually non-existent.

How will the localist react to such examples? That depends on which school he is affiliated with. If he is a noncist, he may well shrug off the evidence and say that, apparently, the contexts in which sentences like these are likely to be used don't call for scalar inferences. This is a lame defence, obviously. We would like to know why scalar expressions evoke upper-bounding inferences in some contexts but not in others, and all the noncist has to offer by way of explanation is that it somehow depends on the context.

For a localist of the defaultist persuasion, like Chierchia or Levinson, the answer has to be that these are cases in which scalar inferences are systematically suppressed. The question then becomes why this is so. Consider example (18a). According to the defaultist, this is preferably understood as, 'I would prefer to visit Tokyo or Kyoto but not both'. The problem is that this isn't an unlikely wish at all, so why should the implicature be cancelled? Or take another example, say (19a). The default interpretation predicted for this sentence is that every farmer owns either goats or sheep but not both. Again this is a perfectly plausible state of affairs, and again it is unclear why the putative local implicature should be cancelled.

Chierchia discusses the following example:

(21)   a.   Everyone wrote a paper or made a presentation. (cf. (19a), (20b))
       b.   Nobody did both.

Chierchia concedes that his theory predicts that (21a) should imply (21b), but argues that this prediction is not wrong:

> I think that the (possible) presence of the strong implicature … is plausible. Suppose we make a bet on (the truth of)
>
> [(21a)]   I bet that everyone wrote a paper or made a classroom presentation.
>
> Then we find out that half of the people did both (while the other half, one of the two). What would happen? I think there would be discussion on who won the bet. If the embedded strong implicature was simply not there, there should be no ground whatsoever for arguing in such a case (Chierchia, 2004, p. 99, n. 6).

This is a poor argument, and a curious one, too. To begin with, unlike Chierchia, I simply don't see that the other party has a leg to stand on, and bets are always followed by post mortem discussion, anyway; so intuitions about this type of scenario are of dubious value, at best. Moreover, coming from a defaultist, this

argument is simply beside the point. Chierchia is committed to the claim that (21a) implies (21b) by default, and therefore his attempt at showing that (21a) may suggest (21b), even if successful, will not further his cause in any substantial way.

Chierchia's theory also predicts that scalar inferences should normally arise in non–monotonic contexts like (19c). Anticipating resistance against this prediction, he writes:

> [(22)]   Exactly two students wrote a term paper or made a classroom presentation.

> Sentence [(22)] can certainly be construed exclusively (as much as it can be construed inclusively). That is, we can intend it to mean that the number of students who did one or the other (but not both) equals two. Or we can intend it to mean that the number of students who did one or the other or both equals two (Chierchia, 2004, p. 47).

Interestingly, Chierchia seems to admit in this passage that exclusive and inclusive interpretations of *or* are equally possible, which goes against his own defaultist doctrine. But apart from this unexpected turn, it is puzzling that Chierchia should suddenly worry about speakers' intentions. For even if his factual claim were correct, which I doubt, it wouldn't begin to prove that the preferred interpretation of (22) is exclusive.

6. Gricean Defence Tactics

The problems that localism is up against are so many and so serious that one may wonder how it ever managed to become such a popular doctrine. On reflection, however, it is clear what happened. The origin of localism lies in a handful of observations that seem problematic for the Gricean approach; these data were taken at face value, and a precipitate diagnosis was mistaken for a general theory of scalar inference.

The obvious thing to do at this juncture is return to the original Gricean account, and consider whether the observations that drove localists to leave the Gricean fold really require a localist explanation. In this section I will have a closer look at the problems summarily introduced in §2, and see what can be done about them within a strictly Gricean framework.

6.1 Disjunction
Consider Chierchia's example (7a) again, repeated here for convenience:

> (23)   Mary is either working at her paper or seeing some of her students.

This sentence presents two problems to the Gricean approach: (i) it has to be explained how (23) could give rise to the implicature that Mary isn't seeing all of her students, without (ii) generating the prediction Mary isn't working at her paper. Sauerland (2004) shows how these problems can be solved on strictly Gricean principles, and matters of detail apart I believe his solution is correct. The critical step in Sauerland's account is to allow that, among the alternatives to a sentence of the form 'φ or ψ' are the sentences φ and ψ as well as *their* alternatives. So the alternatives to (23) include:

(24) a. Mary is working at her paper.
b. Mary is seeing some of her students.
c. Mary is seeing all of her students.

(Note that each of these alternatives is logically stronger than (23).) Applying standard Gricean reasoning, we now infer, for each of these alternatives, that the speaker doesn't know if it is true:

(25) a. $\neg B_s(24a)$
b. $\neg B_s(24b)$
c. $\neg B_s(24c)$

Now suppose the competence assumption holds for (24c): $B_s(24c) \vee B_s\neg(24c)$. Combined with (25c), this assumption entails that $B_s\neg(24c)$, i.e. the speaker assumes that Mary is not seeing all of her students. Thus the first problem is solved.

The implicature in (25c) is strengthened on the basis of the competence assumption. What about the implicatures in (25a,b)? Shouldn't they be strengthened, too? Consider (25a) ((25b) is analogous). If the speaker were competent with respect to (24a), i.e. $B_s(24a) \vee B_s\neg(24a)$, then it would follow that $B_s\neg(24a)$, which would be wrong, of course. However, we cannot make the competence assumption in this case, the reason being that, by uttering (23), the speaker conveys that he doesn't know whether or not (24a) is true. And this follows from the Gricean analysis, too. To see how, consider the case in more abstract terms. The speaker utters a sentence of the form 'φ or ψ', from which we derive $\neg B_s\varphi$ and $\neg B_s\psi$. Assuming, furthermore, that the Quality maxim is honoured, we have $B_s(\varphi \vee \psi)$. Then it follows that $\neg B_s\neg\varphi$ and that $\neg B_s\neg\psi$. In other words: on the assumption that the Quality maxim is not violated, Gricean reasoning allows us to infer, from an utterance of 'φ or ψ', that as far as the speaker knows, φ is possible, $\neg\varphi$ is possible, ψ is possible, and $\neg\psi$ is possible.

Returning to (23): We have just seen how we can infer from an utterance of this sentence that $\neg B_s(24a)$ and that $\neg B_s\neg(24a)$, or in other words: the speaker considers it possible that Mary is working on her paper and he considers it possible that she isn't. Or in yet other words: the competence assumption is not valid in this case. Hence, the implicature in (25a) will not

be strengthened, and the second problem is solved, too—or rather: it doesn't even arise.

6.2 Belief Reports

The following example is Russell's (2006):

(26) a. George believes that some of his advisors are crooks.
      b. George believes that not all of his advisors are crooks.

Intuitively, it seems plausible that an utterance of (26a) licenses the inference that, according to the speaker, (26b) holds, as well, and experimental data confirm this intuition (see §5.4). As discussed in §2, it would seem that Gricean theories fail to account for this observation. The best they can do, apparently, is predict that (26a) implicates that it isn't the case that George believes that all of his advisors are crooks; which isn't wrong, but weaker than (26b).

Several authors have shown how these data can be reconciled with the Gricean view, after all (van Rooij and Schulz, 2004; Spector, 2006; Russell, 2006). These various proposals don't exclude each other, and I believe that more than one of them might be correct. Here I will briefly discuss two ideas that seem especially promising to me. Russell and van Rooij and Schulz propose to adapt the standard Gricean account of scalar inference to belief reports, along the following lines. As we have just seen, the strong scalar implicature which Gricean reasoning will deliver for (26a) is that

(27) $\neg B_g$(all of g's advisors are crooks)

Now it seems plausible to assume that George has an opinion as to whether all of his advisors are crooks. That is to say, we may assume that the competence assumption holds not only for the speaker, but for the subject of the belief report, as well:

(28) B_g(all of g's advisors are crooks) $\vee$ $B_g\neg$(all of g's advisors are crooks)

Between them, (27) and (28) entail (26b), and thus the seemingly local inference is accounted for without resorting to localist trickery.

Spector obtains the same inference in a different way. To explain how, let us first consider the following variation on (26a):

(29) a. George said that some of his advisors are crooks.
      b. Some of my advisors are crooks.

Intuitively, (29a) may suggest that, when he uttered (29b), George believed that not all of his advisors are crooks. Why? Obviously, because it may be assumed that

George's utterance of (29b) implicated so. Armed with this observation, we return to the belief report in (26a), and ask ourselves what sort of evidence the speaker might have for making this claim. In principle, the speaker may be relying on all kinds of sources, but there is one in particular that readily comes to mind: the speaker may know (or at least believe) that (29a) is true. But we have just seen how this proposition might prompt the inference that (26b) is the case, so if an occurrence of (26a) suggests (29a), it may also suggest (26b).

Spector's analysis is less economical than the one proposed by Russell and van Rooij and Schulz. For the latter theory to work, we merely have to assume that the competence assumption holds for the subject of the attitude report, i.e. George. This may not be a plausible assumption in general, but in the case at hand it doesn't seem far-fetched. Spector's story is less straightforward: in order to explain how (26a) comes to imply (26b), we have to assume not only that George said that some of his advisors are crooks, but also that his utterance licensed the implicature that not all of his advisors are crooks. Still, it seems rather likely that, in at least some cases, hearers consider what may be the evidence on which the speaker bases his statement, and if the statement happens to be a belief report, Spector's explanation strikes me as entirely plausible.

No matter which account is the right one, we will have an explanation of the seemingly local scalar inferences associated with belief reports, and one that is fully consistent with Gricean principles. Moreover, whichever explanation we favour, the proliferation problem that besets the localist approach will not arise for our account. As we saw in §5.4, localism predicts that scalar inferences should generally be available within the scope of attitude verbs, while in reality this inference pattern is associated only with *believe* and its synonyms. If we adopt either or both of the proposals discussed above, this problem doesn't arise.

6.3 Presupposition

One of the problems raised by Chierchia in his discussion of the Gricean standard theory concerns factive verbs:

(30) a. Bonnie knows that Clyde took some of the apples.
      b. Clyde didn't take all of the apples.
      c. Clyde took at least some of the apples

An utterance of (30a) may be taken to suggest that, according to the speaker, (30b) holds as well, and as the sentence presupposes (30c), it is natural to suppose that the upper-bounding inference is somehow associated with the presupposition. That is what Chierchia assumes, too: on his account, the implicature originates in the lexical entry for *some*, and thus is already part of the meaning of the complement clause of *know* before the presupposition is triggered. This analysis delivers passable predictions for (30a) (though see Russell, 2006), but not for other factive verbs:

(31)   Bonnie regrets that Clyde took some of the apples.

On its most plausible reading, (31) means that Bonnie regrets that (30c), and conveys, furthermore, that (30b). On Chierchia's localist analysis, however, the presupposition has to be the same as the object of the attitude verb: the presupposition of (31) cannot be different from what Bonnie regrets. Chierchia predicts that, by default, the sentence presupposes (30b) and means that Bonnie regrets (30b); this seems an unlikely reading to me, but that is as it may be. If on the other hand the implicature is cancelled, the sentence ends up presupposing (30c) and meaning that Bonnie regrets (30c). Either way, the presupposition is the same as (hence, not stronger than) the object of Bonnie's attitude.

The trouble with Chierchia's analysis is that his localist convictions lead him to assume that the scalar inferences we observe in (30a) and (31) are part of the presupposed content, to begin with. The obvious alternative is that these inferences derive from what these sentences presuppose. Reverting to the orthodox position that these sentences presuppose (30c), we apply the standard Gricean reasoning to this proposition, and thus arrive at the implicature that, for all the speaker knows, (30b) holds, as well. That (30a) may be taken to convey that Bonnie believes the same thing is explained by the theories discussed in §6.2.

That presuppositions can give rise to scalar implicatures is confirmed by the following examples:

(32)   a.   It was Clyde who took some of the apples.
       b.   It wasn't Clyde who took some of the apples.
       c.   Wasn't it Clyde who took some of the apples?

(33)   a.   The man who took some of the apples will be arrested.
       b.   It's possible that the man who took some of the apples will be arrested.
       c.   How likely is it that the man who took some of the apples will be arrested?

Ceteris paribus, the sentences in (32) presuppose that somebody took some of the apples (in the first case this is an entailment, as well), and according to my intuitions they will also tend to imply that the person in question didn't take all of the apples. Analogous observations apply to the sentences in (33). But does it make sense at all that presuppositions should give rise to implicatures? It does. Consider (30a) again. On all accounts of presupposition I am conversant with, even such which hold that presuppositions are always given *de jure*, (30a) may be used to convey, as information that is new to the hearer, that Clyde took some of the apples. Speaking more generally, it is widely accepted that presuppositional devices are routinely used for introducing new information into the discourse. In this sense presupposition is no different from assertion (which of course is not to deny that they are different

in other respects), and there is every reason to expect that presuppositions should give rise to implicatures. For further discussion of this theme, see Geurts, 2006b.

6.4 Indefinites

Another problem raised by Chierchia involves scalar expressions occurring in the scope of indefinites (see §2). Simplifying and at the same time sharpening Chierchia's example somewhat, suppose I have a story to tell about Clyde, and I start as follows:

> (34)   A friend of mine had an argument with some of his relatives.

Instead of uttering (34), I could have made a stronger statement, namely:

> (35)   A friend of mine had an argument with all of his relatives.

Applying the standard Gricean reasoning, my audience should be entitled to infer, first, that I don't know if (35) is true, and then, that I know that (35) is not true, which is to say that I have no friends who had an argument with all of their relatives. But this is patently absurd: my utterance was about Clyde, and any other friends I may have are so much as irrelevant to my discourse.

What went wrong? The derivation of the fake implicature started out from the question why Clyde didn't say (35) rather than (34). That was the wrong question. What we should have asked is this: 'Why did the speaker say that *the friend in question* had an argument with some of his relatives?' If we answer this question, in the usual Gricean way, we arrive at the implicature that the friend in question (i.e. as it happens, Clyde) didn't have an argument with all of his relatives; which is correct.

Chierchia's problem arises because he is taking it for granted that alternatives enter the Gricean story in a particular way: Chierchia assumes without argument that alternatives are propositions which would have been expressed by other statements the speaker could have made, and that a scalar inference is obtained by negating such a proposition. This may seem like an innocuous assumption to make, but it is the source of our present trouble, and it is not made by Grice himself:

> Anyone who uses a sentence of the form *X is meeting a woman this evening* would normally implicate that *the person to be met* was someone other than X's wife, mother, sister, or perhaps even close platonic friend (Grice, 1989, p. 37; emphasis added).

If Grice had asked why the speaker didn't say that X is meeting *a* close platonic friend, the implicature would have been that X is not meeting any of his close platonic friends. But that isn't how Grice frames the question, and rightly so, because he would have ended up with the wrong implicatures. The point, I

believe, is this. By uttering sentence (34) I did more than merely express a proposition: my utterance introduced a new discourse referent. As this discourse referent (call it x) is available by the time the hearer asks himself why I said what I said, it is natural to assume that this question will be about the same x. Hence, the question is why I didn't say that x had an argument with all of his relatives, and the hearer's answer to this question may well be that I don't know if x had an argument with all of his relatives, or even that I know that he didn't. In any case, the problem noted by Chierchia doesn't arise. Therefore, the root of Chierchia's problem is that it ignores the discourse effects of indefinites.

An unexpected fringe benefit of this analysis, which is developed at greater length in Geurts, 2006b, is that it offers a natural explanation for so-called 'free choice' readings:

(36)   The thief may have been a painter or a composer.

A well-known problem with this type of sentence is that it may be construed as implying that both disjuncts may be true (though not perhaps together). This is nicely accounted for on the assumption that the interpretation of (36) involves a discourse referent W representing the set of possible worlds in which the thief is a painter or a composer. Then the hearer can reason as follows. Could it be the case that the thief is a painter in none of the worlds in W? Presumably not, because then the speaker should have said, 'The thief may have been a composer.' Hence, the thief may have been a painter. When applied to the second disjunct, the same reasoning yields the conclusion that the thief may have been a composer. Thus, free-choice inferences come out as a garden variety of conversational implicature.

7. What the Gricean Theory Can't Do

In §2, I introduced a distinction between two types of facts that, prima facie, suggest that scalar inferences may arise locally, within the scope of operators of various kinds. In the meantime, we have dealt with all C-type examples in ways that are compatible with globalism à la Grice: in all the cases used by Chierchia to argue against globalism, seemingly local implicatures turned out to be amenable to Gricean analysis, after all. Hence, C-type examples may be treated as cases of quasi-local implicature, and give no cause for abandoning the Gricean programme; on the contrary, they demonstrate once again how fruitful that programme is.

I'm less sanguine about the L-type cases, a selection of which I repeat here for ease of reference:

(37)   a.   Drinking warm coffee is better than drinking hot coffee. (= (3a))

    b.   If it's warm, we'll lie out in the sun. But if it's VERY warm, we'll go inside and sit in front of the air-conditioner. (= (4a))

    c.   Around here, we don't LIKE coffee, we LOVE it. (= (5a))

I believe that, in cases like these, we are forced to admit that scalar terms give rise to local upper-bounding interpretations, which cannot be accounted for in terms of implicature; they are local quasi-implicatures. Since this is a non-existence claim, I cannot hope to prove that it is correct; but I will try to make it plausible that L-type cases are special.

One reason for believing that L-type sentences are different from bona fide specimens of scalar implicature is that they are clearly marked.[10] Another reason is that, in the L-type cases, the scalar inference cannot be separated from the Fregean content of the sentence. For instance, if (38a) is understood as implicating that George believes that not all of his advisors are crooks, we can paraphrase it as (38b):

(38)   a.   George believes that some of his advisors are crooks. (= (26a))

          b.   George believes that some of his advisors are crooks and he believes that not all of his advisors are crooks.

Try this with any of the L-type examples, and you will get stuck; which confirms that in these cases the upper-bounding inference is really local and part of the Fregean content of the sentence.

My third reason for maintaining that L-type examples cannot be handled within a Gricean framework is that, to the best of my knowledge, convincing non-localist analyses are simply not available. The only potential exceptions that I know of are cases like (37c), which might be accounted for by Horn's (1985, 1989) theory of 'metalinguistic negation'. However, as argued by McCawley (1991), Carston (1996), and myself (1998a), among others, Horn's theory is problematic in all sorts of ways, and even Horn's own views seem to have evolved away from it: referring to L-type examples like (37a,b), Horn now writes:

---

[10]   Recanati denies that this is invariably the case. Referring to the example,

    (i) Eating some of the cake is better than eating all of it

   he writes that '[e]ven though the word "some" does not bear focal stress, still a contrast is made (by means of the "better than" construction itself) between "some" on the lefthand-side and "all" on the right-hand-side' (Recanati, 2003, p. 323). This observation seems incorrect to me: I believe (i) requires focal stress on *some*, for precisely the reason given by Recanati.

> … in each of these contexts, it's only when the stronger scalar is reached that the earlier, weaker one is retroactively adjusted to accommodate an upper bound into its semantics, e.g. with *some* being *re*interpreted as expressing (rather than merely communicating) 'some but not all'. This reinterpretation is facilitated by the obligatory focus on the relevant scalar operators (Horn, 2006, p. 11).

The mechanism Horn resorts to for dealing with L-type cases has been studied at least since the 1970s (e.g. by Nunberg, 1978, 1979; Bosch, 1984; Carston, 2002; Recanati, 2003). It involves readjusting the Fregean content of an expression, as in:

(39)   I have no marble lions in my garden

If I utter (39), my claim isn't vacuously true on the grounds that marble lions don't exist; rather, the reason I have spoken truthfully is that my garden is devoid of marble statues representing lions. Apparently, the lexical meaning of *lion*, as used here, underwent a reconstrual prior to combining with the other expressions in the sentence, and it is clear that such reconstruals are pragmatic in nature, in the sense that they are very much dependent on contextual information and world knowledge. But it is equally apparent that they aren't implicatures. Reconstrual is a local pragmatic process.

  Although examples like (39) are special in the sense that they are essentially ad hoc, they aren't marked in any way. However, one of the characteristics of L-type instances of scalar inference is that they are marked. The reason for this, I suggest, is that in these cases reconstrual results in a special kind of meaning shift, in that the lexical meaning of an expression is narrowed down; e.g. in (38a) the lexical meaning of *warm* is restricted to 'warm but not hot'. Since, furthermore, the problematic scalars in (38) occur in downward-entailing positions, they will not give rise to implicatures, and therefore it requires effort, on the part of the speaker as well as the hearer, to achieve an upper-bounded interpretation. This is accomplished, in effect, by juxtaposing the scalar term with a stronger scale mate, in such a way that their standard meanings will result in contradiction. Whence the mandatory contrastive stress observed by Horn.

  It bears emphasising that reconstrual is pervasive in everyday language use, as is shown by Nunberg's (1978) pioneering work, and even if the underlying mechanisms are poorly understood, there can be no doubt that it is happening all the time. Therefore, as it seems inevitable that L-type cases of scalar inference should be analysed in localist terms, reconstrual offers itself as the most natural means for carrying through such an analysis.

8. Going Syntactic

The general picture outlined in the foregoing is pragmatic through and through. In the main, scalar inferences are a species of conversational implicature, and in the

exceptional cases where they are truly local, they are due to context-dependent meaning shifts. This view is in stark contrast with grammatical theories developed recently by Chierchia (2006) and Fox (2007), who propose to enrich syntax with an operator whose meaning is more or less the same as that of overt *only*. That such a move isn't entirely devoid of intuitive appeal is shown by Levinson's (1983) observation that some scalar implicatures can be made explicit by means of *only*, as in the following example:

(40)   a.   Clyde stole some of the pears.
         b.   Clyde stole only some of the pears.

When construed as implying that Clyde didn't steal all of the pears, (40a) can be paraphrased as (40b), where *only* associates with the plural indefinite. The key idea underlying the syntax-driven programme is that, whenever (40a) implies that Clyde didn't steal all of the pears, its hidden syntactic form resembles that of (40b) in that (40a), too, contains a form of *only*, albeit a mute one, which I will write MONLY. This covert operator is thought to be freely insertable in the parse tree, thus generating syntactic ambiguities like the following:

(41)   a.   You may have an apple or a pear. (= (16a))
         b.   You MONLY [may have an apple or a pear].
         c.   You may [MONLY have an apple or a pear].

(41a) may be construed with wide scope for covert MONLY, as in (41b), or it may be outscoped by the modal, as shown in (41c). Assuming that the alternatives taken into account by MONLY are generated from the *or*-scale, (41b) rules out that the addressee may have an apple and a pear, while (41c) doesn't.

   Since I don't have the space to discuss either Chierchia's or Fox's theory in any detail, the foregoing sketch completely fails to do justice to the intricacies of the theories they propose. However, even this bare-bones version will suffice for explaining what I take to be the main problems with syntax-driven approaches. To begin with, it should be noted that free insertion of MONLY creates a selection problem of formidable proportions. On the syntax-driven account, every sentence that contains a scalar expression has at least two readings, and the number of possible readings will explode when we start adding further scalar terms and/or scope-bearing expressions. For example, according to Fox, one of the more natural readings of (41a) is to be derived from the following syntactic analysis:

(41)   d.   MONLY MONLY [you may have MONLY an apple or MONLY a pear].

(In Fox's theory, (41d) secures the reading on which the speaker may have an apple or a pear or both.) It will be clear that any syntax-driven theory of scalar 'implicature' stands in need of a module for selecting among the various, and often

numerous, readings predicted to be available for sentences containing scalar expressions. To the best of my knowledge, such a module is still a pie in the sky.

There are more reasons for believing that the syntax-driven approach is on the wrong track, some of which carry over from the discussion of localism in §§4–5. But the biggest reason of them all is that syntactic theories of implicature are blatantly ad hoc. Relegating pragmatic matters to syntax is always a measure of despair, and you need *very* good arguments for doing so. In the foregoing I argued that the whole range of interpretations scalar expressions give rise to is covered quite nicely by two pragmatic mechanisms that are amply motivated on independent grounds: conversational implicature and reconstrual. If this is right, the quest for a syntactic theory of some upper-bounding construals is just pointless.

The opposition between pragmatic and syntactic theories of implicature is not unlike that between Darwinism and creationism, though it has to be said that the syntax-driven approach compares rather favourably to creationism, as it is practised with far greater technical skill and proportionally less religious zeal.

9. Conclusion

As noted in §3, hybrid theories of scalar inference are so common that my main conclusion—that the interpretation of scalar terms is constrained by two very different mechanisms—is unlikely to raise eyebrows. Nor should it be too controversial to say that one of the mechanisms in question is conversational implicature. This much is agreed by most authors. The other conclusions I would like to draw are more contentious:

- Scalar inferences aren't defaults, either in the strong sense that they are automatic, or in the weaker sense that they are made ceteris paribus.
- The local mechanism that occasionally constrains the interpretation of a scalar term is what I have called 'reconstrual'. The argument for this is simply that reconstrual is solidly motivated on independent grounds, and perfectly suitable for this particular job.
- The most natural division of labour between global and local processes of scalar inference is that, as a rule, it is the former that apply, while the latter come into play only in special cases.

The last claim is the opposite from what is held by Bach (1994) and Noveck and Sperber (2007), who maintain that normally speaking scalar inferences result from what I have called 'reconstrual'. It has to be said, however, that these authors don't argue their point; they simply state that this is how things are.

The evidence adduced in the foregoing indicates that local scalar inferences are confined to marked cases. If this constraint is relinquished, we need an alternative account of when and why such inferences arise. There is no such account, and it is unlikely that one can be given. The problem is that, when it is taken beyond the

marked cases, the localist account (any localist account) is bound to fail in many situations, and that the pattern of failure will be quite erratic, and therefore likely to defy explanation. Let me give a few examples to drive home this point. Although the inference in (42a) admits of a localist explanation, we have seen that the one in (42b) does not (§5.2):

(42) a. She ate some of the apples $\rightsquigarrow$ She didn't eat all of them.
      b. She may have eaten some of the apples $\rightsquigarrow$ She will not have eaten all of them. (= (16c))

The Gricean account explains both inferences, and it also explains why the following (= (19)) do not occur:

(43) a. All the farmers in this region own goats or sheep $\not\rightsquigarrow$ None of them own both.
      b. At least 300 of the farmers in this region own goats or sheep $\not\rightsquigarrow$ At least 300 of them don't own both.
      c. There are exactly 300 farmers in this region who own goats or sheep $\not\rightsquigarrow$ There are exactly 300 farmers in this region who don't own both.

As far as I'm aware, there is no localist explanation for this pattern (§5.4). Finally, while the Gricean paradigm allows us to say that (44a–c) are all possible conversational implicatures of (44), localism can only account for the last inference:

(44) Some of my cousins live in Belgium. (= (15))
      a. The speaker is not at liberty to say whether all his cousins live in Belgium.
      b. The speaker doesn't believe that all his cousins live in Belgium.
      c. The speaker believes that not all his cousins live in Belgium.

Taken together with various other data discussed in §§4–6, these observations lend support to the proposed division of labour between scalar implicature and local pragmatics.

Department of Philosophy
University of Nijmegen

References

Bach, K. 1994: Conversational impliciture. *Mind & Language*, 9: 124–162.
Bartsch, R. 1973: 'Negative transportation' gibt es nicht. *Linguistische Berichte*, 27, 1–7.

Begg, I. and Harris, G. 1982: On the interpretation of syllogisms. *Journal of Verbal Learning and Verbal Behavior*, 21, 595–620.

Bezuidenhout, A. 2002: Generalized conversational implicatures and default pragmatic inferences. In J. K. Campbell, M. M. O 'Rourke, and D. Shier (eds), *Meaning and Truth: Investigations in Philosophical Semantics*. New York: Seven Bridges Press, pp. 257–283.

Bosch, P. 1984: *Lexical Learning, Context Dependence, and Metaphor*. Bloomington, IN: Indiana University Linguistics Club.

Bott, L. and Noveck, I. A. 2004: Some utterances are underinformative: the onset and time course of scalar inferences. *Journal of Memory and Language*, 51, 437–457.

Breheny, R., Katsos, N. and Williams, J. 2006: Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition*, 100, 434–463.

Carston, R. 1996: Metalinguistic negation and echoic use. *Journal of Pragmatics*, 25, 309–330.

Carston, R. 1998: Informativeness, relevance and scalar implicature. In R. Carston and S. Uchida (eds), *Relevance Theory: Applications and Implications*. Amsterdam: Benjamins, pp. 179–236.

Carston, R. 2002: *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

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

Chierchia, G. 2006: Broaden your views: implicatures of domain widening and the 'logicality' of language. *Linguistic Inquiry*, 37, 535–590.

Cohen, L. J. 1971: The logical particles of natural language. In Y. Bar-Hillel (ed.), *Pragmatics of Natural Language*. Dordrecht: Reidel, pp. 50–68.

Evans, J. S. and Newstead, S. E. 1980: A study of disjunctive reasoning. *Psychological Research*, 41, 373–388.

Fox, D. 2007: Free choice and the theory of scalar implicatures. In U. Sauerland and P. Stateva (eds), *Presupposition and Implicature in Compositional Semantics*. Basingstoke: Palgrave Macmillan, pp. 71–120.

Geurts, B. 1998a: The mechanisms of denial. *Language*, 74, 274–307.

Geurts, B. 1998b: Scalars. In P. Ludewig and B. Geurts (eds), *Lexikalische Semantik aus kognitiver Sicht. Tübingen: Gunter Narr*, pp. 85–117.

Geurts, B. 2006a: Exclusive disjunction without implicatures. Ms., University of Nijmegen.

Geurts, B. 2006b: Implicature as a discourse phenomenon. In E. Puig-Waldmüller (ed.), *Proceedings of Sinn und Bedeutung* 11. Barcelona: Universitat Pompeu Fabra, pp. 261–275.

Geurts, B. 2006c: Take 'five': the meaning and use of a number word. In S. Vogeleer and L. Tasmowski (eds), *Non-definiteness and Plurality*. Amsterdam/Philadelphia: John Benjamins, pp. 311–329.

Geurts, B. and Pouscoulous, N. 2008: Local implicatures? Forthcoming in: P. Egré and G. Magri (eds), *Presuppositions and Implicatures*. MIT Working Papers in Linguistics.

Grice, H. P. 1989: *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Horn, L. R. 1985: Metalinguistic negation and pragmatic ambiguity. *Language*, 61, 121–174.

Horn, L. R. 1989: *A Natural History of Negation*. Chicago: Chicago University Press.

Horn, L. R. 2006: The border wars: a neo-Gricean perspective. In K. von Heusinger and K. Turner (eds), *Where Semantics Meets Pragmatics*. Amsterdam: Elsevier, pp. 21–48.

Landman, F. 1998: Plurals and maximalization. In S. Rothstein (ed.), *Events and Grammar*. Dordrecht: Kluwer, pp. 237–271.

Levinson, S. C. 1983: *Pragmatics*. Cambridge: Cambridge University Press.

Levinson, S. C. 2000: *Presumptive Meanings*. Cambridge, MA: MIT Press.

McCawley, J. D. 1991: Contrastive negation and metalinguistic negation. In *Papers from the 27th regional meeting of the Chicago Linguistics Society. Part 2: The parasession on Negation*, pp. 189–206.

Newstead, S. E. and Griggs, R. 1983: Drawing inferences from quantified statements: a study of the square of oppositions. *Journal of Verbal Learning and Verbal Behavior*, 22, 535–546.

Noveck, I. 2001: When children are more logical than adults: experimental investigations of scalar implicature. *Cognition*, 78, 165–188.

Noveck, I. and Sperber, D. 2007: The why and how of experimental pragmatics: the case of 'scalar inferences'. In N. Burton-Roberts (ed.), *Pragmatics*. Palgrave Macmillan, pp. 184–212.

Nunberg, G. D. 1978: *The Pragmatics of Reference*. Bloomington, IN: Indiana University Linguistics Club.

Nunberg, G. D. 1979: The non-uniqueness of semantic solutions: polysemy. *Linguistics and Philosophy*, 2, 143–184.

Paris, S. G. 1973: Comprehension of language connectives and propositional logical relationships. *Journal of Experimental Child Psychology*, 16, 278–291.

Pouscoulous, N. 2006: *Processing Scalar Inferences*. Doctoral dissertation, EHESS, Paris.

Recanati, F. 2003: Embedded implicatures. *Philosophical Perspectives*, 17, 299–332.

van Rooij, R. and Schulz, K. 2004: Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information*, 13, 491–519.

Russell, B. 2006: Against grammatical computation of scalar implicatures. *Journal of Semantics*, 23, 361–382.

Sauerland, U. 2004: Scalar implicatures in complex sentences. *Linguistics and Philosophy*, 27, 367–391.

Spector, B. 2006: *Aspects de la pragmatique des opérateurs logiques*. Doctoral dissertation, University of Paris VII.

Sperber, D. and Wilson, D. 1995: *Relevance*, 2nd edn. Oxford: Blackwell.