

Against Grammatical Computation of Scalar Implicatures

BENJAMIN RUSSELL
Brown University

Abstract

Recently, several authors have argued that Gricean theories of scalar implicature computation are inadequate, and, as an alternative, one author has proposed a grammatical system for computing scalar implicatures. The present paper provides arguments, counter to the claims of these authors, that Gricean reasoning can account for the implicatures of certain complex sentences and does not generate undesirable implicatures for others. Moreover, it is shown that a putative advantage of grammatical scalar implicature computation, that it informs a theory of intervention in negative polarity item licensing, is spurious. These arguments, plus general conceptual advantages of Gricean theory, lead to the conclusion that scalar implicature computation is not carried out in the grammar.

1 INTRODUCTION

Scalar implicatures, according to a theoretical tradition that goes back to Grice (1975) and is elaborated and formalized by Horn (1972, 1989), Gazdar (1979), Sauerland (2004), Blutner (2004), and van Rooij & Schulz (2004), have been understood as inferences of the following (somewhat abbreviated) form:

- (1) Speaker A said ϕ , and could as easily have said ψ , which is stronger. Because A is co-operative, she makes the strongest true statement possible, so ψ can't be true.

The key to the reasoning in (1), and to Gricean pragmatics generally, is *co-operation*: it is co-operative to be informative, so scalar implicatures are the product of the distinctly extralinguistic behaviour of agents working together towards a common goal. The idea that such inferences may be attributed to co-operation rather than grammar has been crucial for the development of modern semantic theory. Because the literal meaning of linguistic expressions need not provide a complete description of the understood message of utterances, the theory of grammar only needs to assume a bare-bones model-theoretic semantics

that delivers literal meanings which may then be used by co-operative agents to convey more nuanced information.

In recent years, several challenges to this Gricean view of scalar implicatures have been raised. These arguments take two forms: (a) that there are observed scalar-type inferences that the Gricean theory is unable to compute, and (b) that the Gricean theory generates implicatures that are not observed. In particular, it is argued that the Gricean theory makes wrong predictions about the following types of sentence:

- (2) Gricean analysis fails to predict:
 - a. George believes that some of his advisors are crooks.
 \rightsquigarrow George believes that not all of his advisors are crooks.
 (Chierchia 2004)
 - b. George knows that some of his advisors are crooks.
 \rightsquigarrow Not all of his advisors are crooks. (Chierchia 2004)
 - c. It is better to eat some of the cake than it is to eat all of it.
 \approx It is better to eat some but not all of the cake than it is to eat all of it. (Carston 1988, Levinson 2001, Recanati 2003)
- (3) Gricean analysis incorrectly predicts:
 - a. George ate some of the fries or the apple pie.
 \nrightarrow George did not eat all of the fries or the apple pie.
 (This entails George did not eat the apple pie.) (Chierchia 2004)
 - b. John has more than three children.
 \nrightarrow John has exactly four children. (Fox & Hackl to appear)

A number of theories of scalar implicatures, departing in various ways from Grice's theory, have ensued, developed by Carston (1988), Landman (2001), Levinson (2001) and Recanati (2003), among others. Recently, Chierchia (2004) has proposed a particularly radical departure from the Gricean theory: scalar implicatures, he argues, are not the product of rational behaviour between co-operative conversants as in (1), but rather are computed automatically in the grammar by means of special semantic composition rules and lexical scales. The main argument that has been made against a global, Gricean framework and in favour of a local, grammatical computation mechanism is an empirical one: that a Gricean theory does not generate the set of observed scalar implicatures. Because the Gricean theory follows, without significant further stipulations, from the uncontroversial assumption that speakers co-operate, nobody argues on a theoretical basis that a stipulated mechanism for grammatical implicature computation is

superior to Gricean computation.¹ For this reason, the argument for a general Gricean framework in this paper will consist of demonstrations that claims of over- and under-generation by a purely Gricean theory are unfounded. That is, observed implicatures that have been claimed to be outside the scope of Gricean pragmatics will be derived within such a theory, and claims that Gricean reasoning generates undesirable implicatures will be countered.

There is a second, indirect argument against global Gricean implicature computation provided by Chierchia: he develops a theory of the otherwise mysterious phenomenon of intervention in negative polarity item licensing (Linebarger 1987) that depends on the tools of grammatical implicature computation. In Section 3 I show that this theory does not actually make the right predictions about intervention, and suggest that any theory of intervention based on scalar implicatures is likely to fail. So, in the absence of empirical evidence of the Gricean theory's inadequacy, I conclude that the adoption of a more complex, stipulative grammatical system for the computation of scalar implicatures is unjustified.

2 GRICEAN COMPUTATION OF IMPLICATURES OF COMPLEX SENTENCES

2.1 *Apparent embedded implicatures*

Chierchia identifies a range of observed scalar implicatures that are not obviously generated by Gricean principles. The most compelling of these are apparently embedded implicatures of kind in (4a).

- (4) George believes some of his advisors are crooks.
- a. \rightsquigarrow George believes not all of his advisors are crooks.
 - b. \rightsquigarrow It is not the case that George believes all of his advisors are crooks.

Chierchia argues that Gricean reasoning cannot generate the observed implicature in (4a): if scalar implicatures are derived by the kind of reasoning in (1), hearers can only make conclusions about the negation

¹ There is one exception to this: Fox and Hackl (to appear) argue that a global theory relies crucially on stipulated Horn scales. And, since Horn scales are stipulated formal, linguistic objects (with no apparent function except their use in scalar implicature computation), we may just as well stipulate a grammatical mechanism. I argue briefly below in Section 3 that Horn scales are just a way to specify a class of competing utterances for Gricean reasoning in the absence of a principled theory of competition. Further, even if Horn scales must be assumed, it is still more parsimonious to suppose the lexicon is organized into competition classes (something we have considerable psycholinguistic evidence for) than to introduce a new set of rules into the grammar.

of competing utterances, not embedded clauses within those utterances, and so they should only infer (4b). On the other hand, in Chierchia's system, the implicature of *some of his advisors are crooks* is computed to give a strong meaning of *some of his advisors are crooks and not all of his advisors are crooks*, which subsequently composes up to give a strong meaning of *George believes some of his advisors are crooks and not all of his advisors are crooks*. Thus, by grammaticizing the process of scalar implicature computation, Chierchia is able to compute embedded scalar implicatures like (4a).

But notice that (4a) follows from (4b) in every context where George has *some* belief about whether all of his advisors are crooks: if it is not the case that George believes that all of his advisors are crooks, and George has some belief about whether all of his advisors are crooks, then George must believe that not all of his advisors are crooks.² Speakers I've consulted readily agree that 'Either George thinks all his advisors are crooks or he thinks they're not all crooks' has to be true, suggesting that this is a 'default' background assumption. It might follow from the assumption that George has a belief about each of his advisors' criminality; indeed, in a context where this assumption is explicitly denied, the 'embedded' implicature is not generated:

- (5) George has not yet formed an opinion about all of his advisors, but, at this point, he believes some of them are crooks.

Here, the strongest implicature drawn is that it is not the case that George believes all his advisors are crooks; this is because the default context, in which George is opinionated, is explicitly ruled out. The enrichment of weak implicatures with contextual inferences in a global, Gricean framework correctly mirrors the observed fine-grained context-sensitivity of scalar implicatures. And the fact that Gricean inferencing *can* generate ostensibly embedded implicatures as well as weaker global implicatures obviates the need Chierchia tries to establish here for a grammatical system of implicature computation.³

Scalar terms embedded under factive verbs and other elements with presuppositions are a bit more complex.

² This way of deriving a stronger implicature from a weaker one based on assumptions about an agent's beliefs is akin to the strengthening of epistemic weak implicatures (*it is not the case that the speaker believes . . .*) to epistemic strong implicatures (*the speaker believes it is not the case that . . .*). See Section 2.3 below for details.

³ Incidentally, Chierchia's system, which introduces implicatures at each type *t* meaning, could handle this case, canceling the strong implicature generated for the embedded clause, but not the weaker implicature generated at the matrix clause level. The point here is not that Chierchia's system is inadequate for such cases, but that it is not needed.

(6) George knows some of his advisors are crooks.

This is entailed by the alternative:

(7) George knows all of his advisors are crooks.

Chierchia observes that an utterance of (6) implies not just the negation of (7), as predicted by Gricean reasoning, but also the negation of its presupposition:

(8) All of George's advisors are crooks.

Chierchia claims that his system makes the right predictions about such sentences: implicatures are added an expression's meaning by the compositional semantics at each type *t* meaning (extensionalizing). This means that, in building (6), the embedded sentence is computed to have the strong meaning *some of his advisors are crooks and not all of his advisors are crooks*. When this combines with the predicate *know*, this meaning makes two contributions: it enters into the *belief* relation with George, and it becomes presupposed content, to be projected through the semantic composition by whatever means ordinary presuppositions project. At the end of the day, then, (6) has the following content:

- (9) a. Asserted: George believes some of his advisors are crooks and not all are.
 b. Presupposed: Some of George's advisors are crooks and not all are.

This accords well with intuitions, at least initially.

But contexts where the "implicature of the presupposition" is defeated and yet the ordinary scalar implicature still arises pose a serious problem for this analysis. Consider (6) in the following slightly richer context:

(10) The public has long been aware that every last one of George's advisors is a crook. And now (even) George knows that some of his advisors are crooks.

Here, the right theory will predict that the implicature that George believes not all of his advisors are crooks can still go through while the inference that not all of the advisors are crooks does not. In Chierchia's theory, clauses with scalar terms become ambiguous: they may be strengthened or not. In other words, either the embedded implicature is cancelled (and is found neither in the presupposition nor the assertion), or not (and is found in both).⁴ There is no room for the

⁴ Cancellation, in Chierchia's system, is a re-analysis of a sentence whereby mechanisms for adding implicatures are inactive.

intermediate interpretation found in (10) where the presupposition's 'implicature' is cancelled but the assertion's is not. A rule could be formulated to selectively remove implicatures within the presuppositional system, but this is an extra stipulation, and it makes Chierchia's system still more complex.

A Gricean analysis of such inferences must begin with a careful consideration of their status. Perhaps the first thing to notice is that they are relatively weak compared with scalar implicatures. Indeed, many speakers I consulted do not have the intuition that (6) implies that not all the advisors are crooks at all, but only that it is compatible with such a situation. Such sentences are apparently equally felicitous whether or not all of the advisors are crooks—an expression like *in fact* is not needed to cancel that supposition like it is to cancel an ordinary scalar implicature.

- (11) a. George knows that some of his advisors are crooks, and (in fact) they all are.
 b. Some of George's advisors are crooks, and #(in fact) they all are.

This suggests, not surprisingly within a Gricean framework, that the inference that not all the advisors are crooks is not generated by the same mechanisms as ordinary scalar implicatures.

And, since they are apparently distinct from scalar implicatures, it is beyond the scope of this paper to provide a comprehensive Gricean theory of such 'presuppositional implicatures'—I will, however, make a brief speculation. The source of the inference, made by only some speakers, might be the assumption that George is well-informed about the integrity of each of his advisors. (As an anonymous reviewer points out, this assumption would depend on the speaker's using *know*, rather than *believe*: in canonical (non-downward-entailing) environments, *know*, unlike *believe*, presupposes its subject is well-informed about the propositional content of its complement, and it therefore seems reasonable to suppose it implies its subject is well informed about related propositions.) This assumption, plus the implicature that it is not the case that George believes all his advisors are crooks, licenses the inference that not all of George's advisors are crooks.

Interestingly, such inferences seem to project beyond negation (I've used examples with *possible/certain* because of potential positive-polarity effects getting in the way with *some*):

- (12) George doesn't know that failure in Iraq is possible.
implies Failure in Iraq is not certain.

Certainly *know* in such an environment does not presuppose or imply that its subject is well-informed about anything, let alone that George is well-informed about the odds of failure in Iraq. So such sentences will need to be treated some other way—perhaps through Gricean reasoning about presupposed content. It is plausible that Gricean reasoning incorporates such information: if speakers are aware of competing utterances, they should be aware of the presuppositional content of those utterances as well as their assertoric content. In a case like this, a speaker, considering the alternative *George doesn't know that failure in Iraq is certain*, might choose to use a sentence with a weaker presupposition because she does not want a stronger one to be accommodated. Spelling this out, we have:

- (13) Utterance assertion: \neg George believes failure is possible.
 Utterance presupposition: Failure is possible.
 Alternative's assertion: \neg George believes failure is certain.
 Alternative's presupposition: Failure is certain.

Notice that, whereas the alternative's assertion is weaker than the utterance's, its presupposition is stronger. The speaker, then, could have uttered a sentence with a weaker assertion with a stronger presupposition, but chose not to. The reasons for making such a conversational move will depend on context in ways that are too complex to explore in this paper; but we can say, roughly, that whenever the difference in content between an utterance's presupposition and its alternative's presupposition is of more interest than the difference between the corresponding assertions, we predict a 'presuppositional implicature'. This line of argument is not meant to be a full-fledged Gricean analysis of such inferences; nonetheless, it should suggest that such an analysis is possible, and that its account of sensitivity to context would be more explanatory than a grammatical theory.

A related example has been discussed by Carston (1988), and Recanati (2003), who argue that scalar terms like *some* sometimes have *not all* as part of their meaning (what Carston calls explicature).⁵ Specifically, it seems that *some* is interpreted as *some but not all* in the following sentence:

- (14) It is better to eat *some* of the cake than it is to eat *all* of it.

⁵ This is, of course, different from Chierchia's position: in Chierchia's system, a grammatical rule is responsible for scalar implicature generation, so the syntactic environment a scalar term appears in determines whether an implicature is generated. In contrast, Carston's proposed ambiguity means discourse context determines which *some* is being used, so *some* can mean *some and not all* in any environment. In this paper, I'm arguing specifically against doing it by grammatical rules, but the arguments can be taken as counterarguments to non-grammatical non-Griceans like Carston as well.

Many speakers find such sentences felicitous; critics wonder how that can be if *all* entails *some*. King and Stanley (2006) provide a potential answer to this criticism. They note that (14) requires focal stress on *some*, and they develop a semantics of the restriction of quantification over worlds based on focus. That is, the fact that *some* is focused leads to the restriction of the set of worlds considered by *better* to those where *some and not all* of the cake are eaten. This analysis seems to be a reasonable solution to this problem.

But a potentially simpler solution is available. That is, (14) is simply a comparative—standard semantics for comparatives (Cresswell 1976; Schwarzschild & Wilkinson 2001) gives it a meaning roughly paraphrasable as:

- (15) Any cutoff for *good* that makes *to eat all of the cake* count as *good* also makes *to eat some of the cake* count as *good*.

What makes an infinitive like *to eat all of the cake* count as good? Bare infinitives in present tense have been analysed as generics (Chierchia 1984)—that is, a set of typical worlds or situations in which a ‘generic’ individual eats all of the cake. So *to eat all of the cake* counts as good if this set of situations is included in the extension of *good*. This means that (15) is true iff the generic situations in which all of the cake is eaten are in the extension of *good* only if the generic situations in which some of the cake is eaten are in the extension of *good*. Because situations in which some cake is eaten range from worlds where just a crumb is eaten to those where all is eaten, the extension of the generic *to eat some of the cake* will plausibly exclude situations at either end of the scale, including those where not all the cake is eaten. Therefore, if more restrictive extensions for *good* include situations where less cake is eaten, then (14) comes out true, without appeal to scalar implicatures at all. If this analysis is right, genericity should be crucial for the acceptability of sentences like (14). Indeed, this prediction seems to be borne out: similar sentences in the (usually) non-generic past tense are considerably degraded.

- (16) #It was better to eat *some* of the cake than it was to eat *all* of it.

2.2 Downward-entailing operators

It has been widely observed that implicatures associated with weak scalar terms like *some* are not generated when those terms are located

below a downward-entailing (DE) operator⁶ (Horn 1972); Fauconnier 1975; Gazdar 1979):

- (17) If George eats some of his vegetables, he'll get dessert.
 ↗ It is not the case that if George eats all of his vegetables, he'll get dessert.

Further, scalar terms appearing below DE operators have a *reversed* pattern of observed implicatures:

- (18) If George eats all his vegetables, he'll get dessert.
 ↘ It is not the case that if George eats some of his vegetables, he'll get dessert.

As Levinson (2001) notes, this fact is predicted by a Gricean theory. Suppose δ is a DE environment, and σ and ω are equally marked, with $\sigma \subset \omega$. Then, by the definition of *downward-entailing*, $\delta(\omega) \subset \delta(\sigma)$ —that is, a sentence S with semantics $\delta(\sigma)$ is weaker than (entailed by) the alternative where σ is replaced by ω . This means that an utterance of S , which contains a stronger scalar term, implicates the negation of the sentence with the weaker term, a reversal of the usual pattern. So Gricean theory explains the reversal of scalar implicatures by DE operators, and therefore makes the correct predictions about the implicatures of conditionals like (18).

Chierchia's grammatical theory, in contrast, has to stipulate a special rule for downward-entailing operators: they remove the grammatically computed scalar implicatures of expressions they combine with. Because it is a stipulation, this rule explains nothing—it could just as well have been upward-entailing operators that remove implicatures.⁷ Both Chierchia's theory and a Gricean one make the same predictions about scalar terms within the scope of downward-entailing operators: scalar implicature patterns will be reversed; the Gricean theory, however, explains this behaviour, while Chierchia's accounts for it by a special rule.

2.3 *Scalar terms inside disjunction*

Chierchia argues that the Gricean theory makes the wrong predictions about the interaction of multiple scalar terms in a single sentence: the

⁶ I use *if* for these examples—despite the fact (according to most modern theories of conditionals) that *if* is not itself a downward-entailing operator (see, for example, von Stechow 1999)—since, under the correct circumstances, antecedents of conditionals are downward-entailing *environments*. That is, with a universal modal in the second conjunct, the antecedent is (Strawson) DE, just like the restrictor of a universal quantifier; but with existential modals, the antecedent may be upward-entailing.

⁷ Thanks to Polly Jacobson for suggesting this way of making this point.

reasoning in (1) incorrectly predicts that (19) implicates (19b), not (19a).

- (19) George ate some of the fries or the apple pie.
 a. \rightsquigarrow It is not the case that George ate all of the fries.
 b. $\not\rightarrow$ It is not the case that George ate all of the fries or the apple pie.

The second implicature, which the Gricean theory is supposed to generate, entails that George did not eat the apple pie; this is, of course, undesirable. Following Sauerland (2004), I propose that a reconsideration of the epistemic status of implicatures provides a solution to this problem within Gricean theory.⁸ Sauerland's particular implementation of this solution involves the stipulation of two unrealized lexical items (*L* and *R*) which serve as scalar alternatives to *and* and *or*, as well as the formal apparatus of *crossing scales*. The analysis presented here, though related to Sauerland's, does not rely on stipulated covert lexical items, and is therefore perhaps a clearer demonstration of the applicability of a purely Gricean theory to the complex sentences at issue.

The key to both the analysis presented here and Sauerland's is a closer consideration of the epistemic status of scalar implicatures. Horn (1989) and Soames (1982) have pointed out that the Gricean reasoning for scalar implicature computation is not accurately given in (1); instead, Gricean agents are only justified in drawing the weaker conclusion in (20).

- (20) Speaker A said ϕ , and could as easily have said ψ , which is stronger. Because A is co-operative, she makes the strongest true statement possible, so *A must not know ψ is true* ($\neg K\psi$).

That is, scalar implicatures are inferences of the form $\phi \rightsquigarrow \neg K\psi$, rather than the stronger $\phi \rightsquigarrow K\neg\psi$. Given these weaker inferences, corresponding strong inferences of $K\neg\psi$ may be generated when the hearer assumes that the speaker knows *whether* ψ is true ($K\psi \vee K\neg\psi$)—i.e. when the speaker is *competent* with respect to the truth of ψ . Treating scalar implicatures as this kind of two-stage inference is not only more faithful to Grice's theory; it also provides a solution to Chierchia's puzzle, for (19) now has the implicature in (21), which does not have the bad entailments of (19b).

⁸ A related proposal can be found in van Rooij & Schulz (2004), who present a theory that depends on relevant alternative meanings rather than competing utterances for Gricean reasoning. This is a potentially innocuous departure from classical Gricean reasoning, and I cannot address its consequences here.

(21) $\neg K$ George ate all of the fries or the apple pie.

And (21) entails the following two epistemic facts about the speaker.

- (22) a. $\neg K$ George ate all of the fries.
 b. $\neg K$ George ate the apple pie.

In contexts where the hearer can assume the speaker is competent—i.e. knows whether George ate all the fries—(22a) can be strengthened, yielding the desired implicature in (19a) (in all other contexts, (22a) is predicted to be the strongest implicature drawn, which is very much in keeping with my intuitions).

Why isn't (22b) strengthened to the obviously undesirable inference (23)?

(23) $K\neg$ George ate the apple pie.

This can be explained in Gricean terms: a sentence's scalar implicature cannot be strengthened if this leads to contradiction with another of its basic implicatures. First, notice that p and q are stronger than (entail) $p \vee q$, and p and q are each certainly easier to say than p or q , so if a speaker has chosen to use p or q , she must not know that either p or q is true. So the weak implicatures in (24) should be generated along with the weak implicatures in (21).⁹

- (24) a. $\neg K$ George ate some of the fries.
 b. $\neg K$ George ate the apple pie.

These implicatures, combined with the uncontroversial assumption that speakers believe what they say (25a) (Grice's maxim of quality), are enough to "block" the strengthening of (22b) to (23). The derivation is as follows:

- (25) a. K (George ate some of the fries \vee George ate the apple pie) (from (19))
 b. K (\neg George ate the apple pie \rightarrow George ate some of the fries) (equivalent to (25a))
 c. $K\neg$ George ate the apple pie \rightarrow K George ate some of the fries (from (25b), assuming a distribution axiom for knowledge modality)
 d. $\neg K$ George ate some of the fries \rightarrow $\neg K\neg$ George ate the apple pie (contrapositive of (25c))
 e. $\neg K$ George ate some of the fries (from (24a))
 f. $\neg K\neg$ George ate the apple pie (modus ponens)

⁹ These implicatures are labeled *clausal*, rather than scalar, implicatures by Gazdar (1979).

The conclusion in (25f) contradicts, and therefore blocks, (23), so the observed implicature (19a) is generated by the Gricean theory without any undesirable consequences. The Gricean theory is again capable of generating the right implicatures for complex sentences, counter to Chierchia's argument that the observed implicatures of sentences like (4) and (19) are not generated by a global theory.

2.4 More than *n* phrases

Fox and Hackl (to appear), who develop their own grammatical scalar implicature computation mechanism, discuss data that initially seem problematic for the Gricean theory. Consider the following sentence:

(26) Bill has more than four kids.

Gricean reasoning, Fox and Hackl claim, predicts that this should implicate

(27) Bill has exactly five kids.

Why? Because (28) is a stronger scalar alternative to (26), and it's probably equally marked.

(28) Bill has more than five kids.

Therefore, (28) should be inferred to be false, yielding the interpretation in (27).

This example, however, is amenable to the Gricean treatment given to Chierchia's disjunctive case in the previous section. That is, Gricean reasoning, in fact, only licenses the epistemic weak inference:

(29) \neg K Bill has more than five kids.

And it is reasonable to suppose that there is another highly salient competing utterance for (26): namely, (27) itself. This means a basic scalar implicature of (26) is:

(30) \neg K Bill has exactly five kids.

Because of this, (29) cannot be strengthened, as this would contradict (30). This leaves the intuitively correct implicature that all the speaker knows is that Bill has more than four kids. Again, the Gricean theory, carefully applied, does not make the bad predictions that critics claim it does.

3 A CAVEAT: COMPETING UTTERANCES

It is crucial in the reasoning in (1) and throughout this paper that the hearer and speaker are sensitive to a relation of *easiness*, or *markedness*,

that determines which utterances compete with each other. If this were not the case, scalar implicatures would be generated in great excess. In particular, because every sentence *S* is entailed by an utterance of *S and the sky is blue*, every sentence would implicate that the speaker does not believe the sky is blue. This is clearly not the case. But, because markedness is a key aspect of Gricean reasoning, this fact may be readily explained: for any sentence *S*, *S and the sky is blue* is a lot harder to say than *S*, so neither hearers nor speakers consider *S and the sky is blue* a competing utterance for *S* (except, perhaps, in certain bizarre contexts). Gricean reasoning is limited to stronger *competing* utterances—things a speaker might just as easily have said. It is not within the scope of this paper to provide a worked-out theory of which utterances compete—nonetheless, I assume that such a theory could be developed, and speculate briefly about what such a theory would look like in this section.

A possible (but probably incorrect) theory of competition is a purely phonological one: one expression competes with another iff it has fewer syllables (a theory like this is advanced by McCawley (1978) and refuted in Horn (1978)). A more principled theory of competition has not been widely pursued—Matsumoto (1995) contains the best-developed theory, and even this provides only rudimentary first steps, albeit promising ones. Instead, theories of implicature have taken competition to be determined by stipulated lexical scales (often called Horn scales) such as ⟨some, all⟩ and ⟨possibly, certainly⟩; two utterances compete with each other if their differences are limited to substitution of scalemates—i.e. they are scalar alternates. But this should not be understood to mean (as suggested by Fox and Hackl (to appear)) that Gricean pragmatics depends on the stipulation of lexical scales. From a Gricean perspective, although it is uncontroversial that scalar alternates are equally marked expressions, and therefore can enter into the type of reasoning given in (1), it is also clear that stipulated Horn scales provide just one of many possible theories of competition (and not a particularly interesting or principled one). Furthermore, there is no clear sense in which Horn scales must be *in* the lexicon—rather, they are generalizations *about* the lexicon: certain lexical items compete with others; saying two items are Horn scalemates is just a way of saying they are competitors (with the particular semantic relation of asymmetric entailment), and it is well-known that scalar implicatures may be generated based on contextually-determined competitors (see Hirschberg 1991 on ‘on-the-fly’ construction of Horn scales). Finally, there is psycholinguistic evidence that shows that speakers reason about competing utterances (see, for

example, Sedivy 2003 and Grodner and Sedivy (to appear)). Because of this, assuming competition exists, and therefore that not all stronger utterances are ruled out by Gricean reasoning, is a less costly assumption than the alternative—that there are grammatical mechanisms for scalar implicature computation.

4 SCALAR IMPLICATURES AND INTERVENTION

Ladusaw (1979) made the following seminal proposal: negative polarity items (like *any*, *ever*, and *drink a drop*) are only licensed when they fall within the scope of a downward-entailing operator. Many subsequent theories of NPI licensing have tried to explain (Lee & Horn 1994; Kadmon & Landman 1993; Krifka 1995; Lahiri 1997) and made refinements to (Linebarger 1987; Zwarts 1996; Giannakidou 1998) this generalization. One particular refinement to this generalization (due to Linebarger 1987) has apparently stumped researchers: an NPI is not licensed by a DE operator if it is in the scope of *every* or contained in one of the conjuncts of *and* (where *every* and *and* are also within the scope of the DE operator).

- (31) a. *Condoleeza didn't drink any Coke and a milkshake. (cf. Condoleezadidn't drink any Coke or a milkshake.)
 b. *George didn't believe Condoleeza had ever eaten a McRib and Dick had ordered a McDLT. (cf. ...Condoleeza had ever eaten a McRib or Dick had ordered a McDLT.)
 c. *Dick doubts everyone would lift a finger to help him finish his McNuggets. (cf. Dick doubts George would lift a finger to help him finish his McNuggets.)

In (31), the negative polarity items *any*, *ever*, and *lift a finger* appear within the scope of downward-entailing operators (*doubt* and negation), yet the sentences they appear in are ungrammatical.

Chierchia develops a theory of NPI licensing based on Kadmon & Landman's (1993) semantics of *any* that provides a putative solution to the intervention puzzle. Kadmon & Landman propose that *any* means approximately the same thing as *some* or *a*: it is an existential quantifier. *any*, however, requires a wider domain of quantification than *some*: if *some* quantifies over, say, the set of regular bun-and-beef-patty burgers, *any* quantifies over a wider domain, including McDLTs, veggie burgers, lamburgers, month-old Big Macs in the garbage, and so on. This aspect of *any*'s meaning is called *Widening*. A consequence of Widening is that it makes *any* semantically weaker than *some*, or $[[\text{some}]] \subset [[\text{any}]]$. To Widening, Kadmon and Landman add a

quasi-pragmatic condition called *Strengthening*: *any* is not licensed unless it strengthens the meaning of the sentence it is in. That is, a sentence S_{any} containing *any* must entail the corresponding sentence S_{some} obtained by substituting *some* for *any*. Since $\llbracket some \rrbracket \subset \llbracket any \rrbracket$ (Widening), Strengthening ($\llbracket S_{any} \rrbracket \subset \llbracket S_{some} \rrbracket$) is not satisfied unless *any* finds itself in a downward-entailing environment. So Kadmon & Landman make roughly the same prediction as Ladusaw: *any* only appears in downward entailing environments. And, since the intervention environments found in (31) are downward-entailing, they make the wrong predictions about intervention cases. Before dealing with intervention, however, a few comments about Kadmon and Landman's theory are in order.

Kadmon & Landman's theory is particularly attractive because its two aspects are relatively well-motivated: Widening by intuitions about the meaning of *any* and its use in discourse (see Kadmon & Landman for details), and Strengthening by general pragmatic considerations: what would be the point of using *any* if it didn't lead to a more informative sentence? But, as attractive as this picture is, the Strengthening condition cannot be purely pragmatic in this way. After all, the theory of scalar implicatures provides an obvious answer to the question just raised: the point of a speaker using a weaker term might be that the speaker does not have evidence to support the entailments of the stronger term. For this reason, assuming Widening, uttering a sentence like (32a) should be perfectly grammatical, and it should (in most contexts) have a scalar implicature: that the speaker was not in a position to make the stronger assertion in (32b).

- (32) a. *George ate any McNuggets.
 b. George ate some McNuggets.

So, given Widening, *any* used in upward-entailing environments should be predicted to be (factoring in its implicatures) an existential over fringe cases: (32a) should mean that George did eat McNuggets, if you're lenient about what counts as McNuggets (allowing, say, limited edition McNuggets made from beef parts), but imply he didn't eat normal McNuggets in a normal way. So Strengthening doesn't follow from pragmatic principles plus the weakness of *any*, so if it is the right constraint, it seems it might have to be stated as part of the grammar.

Moreover, Kadmon & Landman recognize that their theory doesn't quite capture Ladusaw's generalization: Ladusaw proposed NPIs were licensed by downward-entailing *operators*; but Strengthening as stated

above predicts they should be licensed in downward-entailing *environments*. The former, not the latter, makes the right predictions about NPIs embedded below two downward-entailing operators (negation and the restrictor of the quantifier *every*): they are licit.

- (33) Donald didn't think that everyone who left any fries on their tray would get fired.

The global environment *any* appears in (33) is upward-entailing. This fact, which Chierchia calls the *roofing* effect, is crucial: it means NPIs are licensed below the root clause level. It also means that any process that affects NPI licensing must also be available below the root clause level.

Chierchia argues that scalar implicatures affect the licensing of NPIs, specifically in intervention cases, and that this fact requires a theory like his that computes scalar implicatures in embedded environments. Chierchia's argument is as follows: *every* and *and* are strong scalar items, with weaker alternatives *some* and *or*, respectively. When these scalar items appear in NPI-licensing downward-entailing environments, the sentences that contain them have scalar implicatures. Chierchia's idea is that these scalar implicatures are what prevent the licensing of *any* in intervention environments; that Kadmon & Landman's Strengthening condition should be strengthened, requiring *any* to strengthen not only the ordinary meanings of their counterparts with *some*, but also their scalar implicatures. Sketchy details of his implementation of this idea are as follows.

Chierchia adapts Kadmon & Landman's theory, introducing into it a modified Strengthening condition that depends on scalar implicature-enriched *strong meanings*. A *strong meaning* is the conjunction of the regular semantics of an expression with its locally-computed scalar implicatures; given an expression α , the strong meaning of α is notated $\llbracket \alpha \rrbracket^s$; to illustrate:

- (34) $\llbracket \text{Condoleeza didn't drink some Coke and a milkshake} \rrbracket^s$
 $\approx \llbracket \text{Condoleeza didn't drink some Coke and a milkshake, but Condoleeza drank some coke or a milkshake} \rrbracket$

Chierchia's idea is that Kadmon & Landman's Strengthening condition is too permissive: it only requires *any* to strengthen ordinary meanings of corresponding *some* sentences. Instead, why not make it strengthen strong meanings, disallowing *any* unless it yields a sentence that is stronger than the corresponding *some* sentence plus its scalar implicatures. Chierchia's Strong Strengthening condition, then, is $\llbracket S_{\text{any}} \rrbracket \subset \llbracket S_{\text{some}} \rrbracket^s$; in prose, the condition is that *any* is not licensed unless the meaning of a (possibly embedded) sentence with *any* is stronger than

(entails) the strong meaning of the corresponding sentence with *some*. Now, in 4, $\llbracket S_{\text{some}} \rrbracket^s \approx \llbracket \text{Condoleeza didn't drink some Coke and a milkshake, but she drank some coke or a milkshake} \rrbracket$, so, in particular, for *any* to be licensed, S_{any} —i.e. (31a)—must entail S_{some} 's scalar implicature, *Condoleeza drank some Coke or a milkshake*. S_{any} does not entail this (not drinking both doesn't entail drinking one), so Strong Strengthening is not satisfied, and (31a) is ungrammatical. Crucially, Strong Strengthening must reference the implicatures of an embedded sentence; so if Strong Strengthening is right, scalar implicatures must be available below the root clause level, indirect evidence for Chierchia's grammatical theory of scalar implicature computation.

But Chierchia's Strong Strengthening condition does not stand up to close scrutiny. Remember that Strong Strengthening requires S_{any} to entail not only S_{some} , but also the scalar implicatures of S_{some} . In Chierchia's system, S_{any} has a grammatically computed scalar implicature as well; this is roughly equivalent to that of S_{some} , except it is weaker—i.e., $\text{imp}(S_{\text{some}}) \subset \text{imp}(S_{\text{any}})$.¹⁰ So, as long as S_{any} doesn't entail its own scalar implicature, it won't entail the implicature of S_{some} . And sentences, in Chierchia's system, don't entail their own implicatures: implicatures are negations of stronger, non-null alternatives. This means that when a DE operator has a strong scalar term and an NPI in its scope, an implicature will be generated, Strong Strengthening will not be satisfied, and the NPI will not be licensed. This all means that, in Chierchia's theory, *and* and *every* are interveners simply because they have scalar implicatures when they are below downward-entailing operators.¹¹

This solution makes two incorrect predictions about intervention. First, every scalar item that has scalar implicatures when embedded below a downward-entailing operator (i.e. every non-weakest scalar

¹⁰ Here's a sketch of a proof: let δ be a downward-entailing environment that a strong scalar term σ (with weaker counterpart ω) and *any* are embedded below. Schematically, then, S_{any} is $\delta(\sigma, \text{any})$, and $\text{imp}(S_{\text{any}}) = \neg\delta(\omega, \text{any})$. S_{some} is, correspondingly, $\delta(\sigma, \text{some})$, and $\text{imp}(S_{\text{some}}) = \neg\delta(\omega, \text{some})$. Now, since $\text{some} \subset \text{any}$, and δ is DE, $\delta(\omega, \text{any}) \subset \delta(\omega, \text{some})$, so $\text{imp}(S_{\text{some}}) = \neg\delta(\omega, \text{some}) \subset \neg\delta(\omega, \text{any}) = \text{imp}(S_{\text{any}})$.

¹¹ As Chierchia formulates the Strong Strengthening condition, ordinary meanings of *any* sentences must entail strong meanings of *some* sentences. But the quasi-pragmatic reasoning given for Strong Strengthening should require the *strong* meaning of an *any* sentence to entail the strong meaning of the corresponding *some* sentence (i.e. $\llbracket S_{\text{any}} \rrbracket^s \subset \llbracket S_{\text{some}} \rrbracket^s$). Apparently, such a condition would not definitively rule out intervention cases. Here are consistent denotations for S_{any} , S_{some} , and their grammatically computed scalar implicatures from (31a) that satisfy this flavour of Strong Strengthening: $\llbracket S_{\text{some}} \rrbracket = \{w_1, w_2, w_3, w_4\}$, $\llbracket S_{\text{any}} \rrbracket = \{w_1, w_3\}$, $\text{imp}(S_{\text{some}}) = \{w_3, w_4, w_5\}$, and $\text{imp}(S_{\text{any}}) = \{w_2, w_3, w_4, w_5\}$, so we have $\llbracket S_{\text{any}} \rrbracket^s = \{w_3\}$ and $\llbracket S_{\text{some}} \rrbracket^s = \{w_3, w_4\}$. The interested reader can check to make sure all necessary entailment relations hold. Given the untenable nature of this alternate Strong Strengthening, only Chierchia's actual formulation of Strong Strengthening will be considered in the rest of the paper.

term) is predicted to be an intervener. So, in addition to *and* and *every*, the set of interveners should include every strong scalar term, like *necessarily*, *required to*, and *excellent*, because no sentence with one of these items below a single downward-entailing operator will satisfy Strong Strengthening. But this prediction is contradicted by empirical data:

- (35) a. I don't think Donald will necessarily eat any cheeseburgers.
 b. Donald doesn't believe George is required to inform anyone before he supersedes his meal.

In (35a), S_{some} implicates *I think Donald will possibly eat some cheeseburgers*; in (35b), S_{some} implicates *Donald believes George is allowed to inform Congress*. Strong Strengthening requires that each corresponding S_{any} —i.e. (35a) and (35b)—entail these implicatures, respectively. But, of course, they do not: these are roughly the implicatures, not entailments, of (35a) and (35b). So strong scalar terms are not, in general, interveners, a fact that is incompatible with the Strong Strengthening theory of NPI licensing. Chierchia notices in his paper that *few*, predicted by Strong Strengthening to intervene, does not; he invokes a difference between primary and indirect implicatures to account for this. But no such distinction is apparent in the scalar terms in (35); indeed, it is hard to imagine a theory of intervention that is based on the strong scalarity of *and* and *every* that does not predict *necessarily* and *required to* also intervene.¹²

The second incorrect prediction made by Chierchia's theory of intervention is that whenever an NPI and a non-weakest scalar term are within the scope of a downward-entailing operator, the scalar term will intervene in the licensing of the NPI, regardless of their relative configuration. This is because a scalar implicature will be generated when the expression containing the scalar term and the NPI combines with the downward-entailing operator, and so this expression will have a non-trivial strong meaning that it must entail. Since such an entailment, i.e. the entailment by a sentence of its own implicature,

¹² Chierchia also discusses numerals, which his theory predicts will intervene unless they are the lowest element of some relevant scale. He cites the following contrast as evidence that this prediction is borne out:

- (36) a. I didn't meet eleven people that read some/*any of my poetry.
 b. I never had eleven kids who won any championship. (said by a soccer coach)

To my ear and my informants', the contrast is not so sharp (*I never had eleven kids who won a championship*, which Chierchia does not mention, is much better than (36b)). Furthermore, the prediction that numerals are interveners depends on a scalar analysis of numerals, an assumption that there is considerable evidence against (see Horn 1992 and references therein for arguments). Because of this evidence, I've left arguments involving numerals out of the present paper for the most part, and won't address in depth any possible intervention facts that involve them.

never goes through, NPIs in such configurations are expected never to be licensed. But they can be perfectly grammatical.¹³

- (39) a. If you ever order every item on the Value Menu, you'll be appointed Secretary of Great Deals.
 b. Colin does not think anyone has eaten a Big Mac and a Quarter Pounder.

In (39b), e.g. S_{some} implicates *Colin thinks someone has eaten a Big Mac or a Quarter Pounder*, which, in turn, is not entailed by (39b), and so Strong Strengthening is not satisfied. A correct theory of intervention should only predict that an NPI is not licensed when it is *in the scope* of an intervener, not whenever an NPI and an intervener are both in the scope of a licenser.

The incorrect characterization of the class of interveners and the scope conditions necessary for intervention is not an accident of Chierchia's way of formalizing Strong Strengthening. It is a consequence of the conceptual basis of the theory. Strong Strengthening is supposed to be an *explanation* of intervention: that it is a consequence of 'the interplay of the general licensing condition on *any* (it must lead to strengthening with respect to some), and the way strong meanings (i.e., implicatures) are computed.' (p. 90) But *any* can't (strongly) strengthen a constituent that has a grammatically-computed scalar implicature, so strong scalar elements, which generally have implicatures in NPI-licensing environments, are incorrectly predicted by Strong Strengthening to always be interveners. Likewise, the presence of grammatically-computed scalar implicatures in a constituent is unaffected by the relative scope of the NPI and scalar term, so Strong Strengthening is, in principle, insensitive to scope. The facts that many strong scalar terms are not interveners and that intervention is sensitive to scope casts doubt on the claim that Strong Strengthening, or anything like it, can explain intervention.

It is not within the scope of this paper to propose an alternate theory of intervention in NPI-licensing. But the facts presented in this section suggest that it is not the scalar implicatures associated with *and* and *every* that are responsible for their behavior as interveners in the

¹³ An anonymous reviewer wonders whether *anyone* in examples like these could be free choice *any*, failing to show intervention effects because it scopes outside negation. But, by the usual diagnostics, free choice *any* is not licensed in these environments:

- (37) *Colin does not think nearly anyone has eaten a Big Mac and a Quarter Pounder.

Moreover, the possibility of the NPI *ever*, which has no free choice counterpart, in this environment provides additional evidence that it is, in fact, an NPI-licensing environment:

- (38) Colin does not think anyone has ever eaten a Big Mac and a Quarter Pounder.

grammatical process of NPI licensing. Indeed, it seems unlikely that a successful theory of intervention in NPI licensing will be based on scalar implicatures, and so intervention does not provide a reason to think that scalar implicatures are grammatically computed.

5 CONCLUSION

Chierchia's proposal to grammaticize the computation of scalar implicatures is a radical departure from Gricean pragmatics. If Chierchia is right, scalar implicatures have nothing to do with rational principles of co-operation in conversation. Instead, this broad class of inferences that can be explained by the simple assumption that speakers co-operate is redundantly generated by a stipulated grammatical mechanism. Arguments for the radical shift to a grammatical theory of implicature computation depend on two claims: the first is that there are observed, apparently scalar inferences the Gricean theory can't generate, as well as undesirable inferences that Gricean reasoning does generate. I've argued that a very general global Gricean theory (compatible with specific formalisms of this theory recently advanced in the literature) makes the right predictions about such inferences. The second argument, recently advanced by Chierchia, is that a semantic theory of intervention in NPI licensing depends on grammatically computed scalar implicatures; I've argued that any theory that depends on scalar implicatures to exclude intervention cases is too stringent, failing to license *any* in many grammatical sentences. I conclude that a departure from Gricean pragmatics is unwarranted, and we can, at least at this stage of the game, keep scalar implicatures out of the grammar.

Acknowledgements

This paper has benefitted from conversations with Polly Jacobson, Chris Barker, Dan Grodner, Julie Sedivy, Larry Horn, Robert van Rooij, Danny Fox, Brian Weatherston, and the audience at the ESSLI 2004 workshop on Implicature and Conversational Meaning. I am especially grateful to Bart Geurts and two anonymous reviewers for their extremely thorough and helpful critiques of earlier drafts of this paper.

BENJAMIN RUSSELL
Brown University
Department of Cognitive and Linguistic Sciences
Providence RI 02912, USA,
e-mail: brussell@brown.edu

First version received: 1.12.05
Second version received: 7.7.06
Accepted: 18.10.06

REFERENCES

- Blutner, R. (2004) 'Pragmatics and the lexicon'. In L. Horn & G. Ward (eds), *Handbook of Pragmatics*. Blackwell.
- Carston, R. (1988) 'Implicature, explication, and truth-theoretic semantics'. In R. Kempson (ed.), *Mental Representations: The Interface Between Language and Reality*. Cambridge University Press. Cambridge, England, 155–181.
- Chierchia, G. (1984) *Topics in the Syntax and Semantics of Infinitives and Gerunds*. PhD thesis, University of Massachusetts.
- Chierchia, G. (2004) 'Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface'. In A. Belletti (ed.), *Structures and Beyond*. Oxford University Press.
- Cresswell, M. J. (1976) 'The semantics of degree'. In B. H. Partee (ed.), *Montague Grammar*. Academic Press. New York, 261–292.
- Fauconnier, G. (1975) 'Pragmatic scales and logical structure'. *Linguistic Inquiry* 6(3):353–375.
- Fox, D. & Hackl, M. (to appear) 'The universal density of measurement'. *Linguistics and Philosophy*.
- Gazdar, G. (1979) *Pragmatics: Implicature, Presupposition, and Logical Form*. Academic Press. New York.
- Giannakidou, A. (1998) *Polarity Sensitivity as (Non) Veridical Dependency*. John Benjamins. Amsterdam.
- Grice, H. P. (1975) 'Logic and conversation'. In D. Davidson and G. Harman (eds), *The Logic of Grammar*. Dickenson Publishing Co. Encino, California, 64–75.
- Grodner, D. & Sedivy, J. (to appear) 'The effect of speaker-specific information on pragmatic inferences'. In N. Pearlmutter and E. Gibson (eds), *The Processing and Acquisition of Reference*. MIT Press. Cambridge.
- Hirschberg, J. (1991) *A Theory of Scalar Implicature*. Garland Publishing Company. New York.
- Horn, L. (1992) 'The said and the unsaid'. *Proceedings from Semantics and Linguistic Theory II*, Cornell University, Ithaca, New York, pp. 163–192. Available from CLC Publications, Department of Linguistics, Morrill Hall, Cornell University, Ithaca, NY 14853-4701.
- Horn, L. R. (1972), *On the Semantic Properties of Logical Operators in English*. PhD thesis, University of California, Los Angeles.
- Horn, L. R. (1978) 'Lexical incorporation, implicature, and the least effort hypothesis'. In D. Farkas (ed.), *Papers from the Parasession on the Lexicon*. Chicago Linguistics Society. Chicago, 196–209.
- Horn, L. R. (1989) *A Natural History of Negation*. Chicago University Press. Chicago, IL.
- Kadmon, N. and Landman, F. (1993) 'Any'. *Linguistics and Philosophy* 16(4):353–422.
- King, J. C. and Stanley, J. (2006) 'Semantics, pragmatics, and the role of semantic content'. In Z. Szabo (ed.), *Semantics vs. Pragmatics*. Oxford University Press.
- Krifka, M. (1995) 'Focus and the interpretation of generic sentences'. In G. Carlson and F. J. Pelletier (eds), *The Generic Book*. Chicago University Press. Chicago, IL, 238–264.
- Ladusaw, W. A. (1979) *Polarity Sensitivity as Inherent Scope Relations*. PhD thesis, University of Texas, Austin.
- Lahiri, U. (1997) 'Focus and negative polarity in Hindi'. *Natural Language Semantics* 6(1):57–123.
- Landman, F. (2001) *Events and Plurality: The Jerusalem Lectures*, Vol. 76 of

- Studies in Linguistics and Philosophy*. Kluwer Academic Publishers.
- Lee, Y.-S. and Horn, L. (1994) *Any* as indefinite plus *even*. Manuscript, Yale University.
- Levinson, S. C. (2001) *Presumptive Meanings*. The MIT Press. Cambridge, MA.
- Linebarger, M. C. (1987) 'Negative polarity and grammatical representation'. *Linguistics and Philosophy* **10**(3):325–387.
- Matsumoto, Y. (1995) 'The conversational condition on Horn scales'. *Linguistics and Philosophy* **18**(1):21–60.
- McCawley, J. D. (1978) 'Conversational implicature and the lexicon'. In P. Cole (ed.), *Syntax and Semantics 9: Pragmatics*. Academic Press. New York, 245–258.
- Recanati, F. (2003) 'Embedded implicatures' *Philosophical Perspectives* **17**(1):299–332.
- Sauerland, U. (2004) 'Scalar implicatures in complex sentences'. *Linguistics and Philosophy* **27**:367–391.
- Schwarzschild, R. and Wilkinson, K. (2001) 'Quantifiers in comparatives: A semantics of degree based on intervals'. *Natural Language Semantics* **10**(1):1–41.
- Sedivy, J. (2003) 'Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations'. *Journal of Psycholinguistic Research* **32**(1):3–23.
- Soames, S. (1982) 'How presuppositions are inherited: A solution to the projection problem'. *Linguistic Inquiry* **13**:483–545.
- van Rooij, R. and Schulz, K. (2004) 'Exhaustive interpretation of complex sentences'. *Journal of Logic, Language and Information* **13**:491–519.
- von Stechow, K. (1999) 'NPI licensing, Strawson entailment, and context dependency'. *Journal of Semantics* **16**(2):97–148.
- Zwarts, F. (1996) 'Three types of polarity'. In F. Hamm and E. Hinrichs (eds), *Plurality and Quantification*. Kluwer. Dordrecht, 177–238.