Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences

Richard Breheny[a,*], Napoleon Katsos[b], John Williams[b,1]

[a]*Department of Linguistics, University College London, London WC1E 6BT, UK*
[b]*RCEAL, University of Cambridge, Cambridge, UK*

Abstract

Recent research in semantics and pragmatics has revived the debate about whether there are two cognitively distinct categories of conversational implicatures: generalised and particularised. Generalised conversational implicatures are so-called because they seem to arise more or less independently of contextual support. Particularised implicatures are more context-bound. The Default view is that generalised implicatures are default inferences and that their computation is relatively autonomous - being computed by some default mechanism and only being open to cancellation at a second stage when contextual assumptions are taken into consideration (Chierchia, 2004; Horn, 1984; Levinson, 2000 i.a.). It is at that second stage where contextual assumptions are considered that *particularised* implications are computed. By contrast, Context-Driven theorists claim that both generalised and particularised implicatures are generated by the same process and only where there is contextual support (Carston, 1998; Sperber & Wilson, 1986 i.a.). In this paper, we present three on-line studies of the prototypical cases of generalised implicatures: the *scalar implicatures* 'some of the Fs' > 'not all the Fs' and 'X or Y' > 'either X or Y but not both'. These studies were designed to test the context-dependence and autonomy of the implicatures. Our results suggest that these scalar implicatures are dependent on the conversational context and that they show

none of the autonomy predicted by the Default view. We conclude with a discussion of the degree to which such implicatures are purely context-driven and whether an interactionist default position may also be plausible.

1. Introduction

Given John's question in (1), Mary's answer would normally imply (2a) and (2b).

1. John: Was the exam easy?
   Mary: Some of the students failed
2a. Not all the students failed
2b. The exam was not easy

It has been suggested that, although not entailments, pragmatic implications like (2a) show a degree of regularity and have the intuitive feel of components of conventional meaning, whereas implications like (2b) may seem to be more context-dependent and indirect. For example, given a different question, 'Is the teacher doing a good job?' implication (2a) should remain, but implication (2b) may not arise and other inferences may become available pertaining to the teacher's skills. Implications like (2a) have been called Generalised Conversational Implicatures and implications like (2b), Particularised Conversational Implicatures. While there is broad agreement that the computation of Particularised Conversational Implicatures (PCIs) would require a general pragmatic system that assesses linguistic input against contextual assumptions based on conversational principles, a central issue in the pragmatics literature concerns the status of the less context-dependent implicatures. The question is whether computing Generalised Conversational Implicatures (GCIs) involves an additional, distinct process which is encapsulated or based on default rules. The additional process would generate GCIs. As defeasible conversational implicatures they would then be open to cancellation by the general pragmatic system should discourse goals require so (Gazdar, 1979; Horn, 1982).

In this paper, we consider two recent versions of the two-system, Default view. According to Chierchia (2004), GCIs are computed as part of the grammar and, according to Levinson (2000), GCIs are computed by a specialised pragmatic system that yields generalised inferences. In both cases, GCIs may have to be cancelled at a later stage when contextual assumptions are taken into consideration. Therefore, although GCIs are computed by a system of default rules, they are 'defeasible' inferences-like PCIs. So, in both cases, one system deals with GCIs (which are linguistically mandated) and a separate one deals with 'genuine' pragmatic implications like PCIs. Both Chierchia's and Levinson's accounts explicitly predict that the 'defaultness' of GCIs applies at the level of on-line language processing.

In contrast to the Default view, there is a view which does not postulate different types of rules for GCIs and PCIs. This Context-Driven approach holds that both types of

implicatures are computed by a single pragmatic system which takes contextual considerations into account from the first stage (Carston, 1998; Hobbs et al, 1993; Sperber & Wilson, 1986). In this approach all pragmatic implications are primarily dealt with outside the grammar by general pragmatic inference. Again, psycholinguistic validation has been called for (Wilson & Sperber, 2003).

There is a small, but growing number of studies of the on-line processing of pragmatic implicatures in text comprehension and reasoning. Single sentence truth value judgment tasks by Noveck and Posada (2003); Bott and Noveck (2004) consistently support the Context-Driven view. However, results of text comprehension studies by Bezuidenhout and Cutting (2002); Bezuidenhout and Morris (2004) have been equivocal.

In this paper, we present a text comprehension experiment in which we attempted to replicate Bezuidenhout and Cutting's study in a simplified form - involving disjunction. We found the results to favour the Context-Driven view, thus falling in line with the single sentence truth value judgment tasks. In addition, we present two experiments where we tested whether scalar implicatures arising from the partitive phrase, 'some of the Fs' are computed in genuinely neutral contexts. The findings of this new type of study were also in accord with the predictions of the Context-Driven approach, disconfirming the Default view.

While there has been some debate already about the generation of common conversational implicatures, it has yet to be located in the broader context of research on language processing. We would argue that the question of at what stage contextual assumptions affect implicature generation is related to research done on sentence comprehension. In particular it is related to the question of whether the human parser is primarily structure based, relying on certain syntactic criteria (Ferreira & Clifton, 1986; Frazier & Clifton, 1996; Frazier & Fodor, 1978), or whether different kinds of information (e.g. from syntax, semantic plausibility, the linguistic context, the situation referred to, frequency based) interact from the earliest possible stage (Altmann & Steedman, 1988; Spivey-Knowlton & Sedivy, 1995; Taraban & McClelland, 1988; Trueswell, 1996).

We conclude this paper with a discussion of our findings in the context of these issues — suggesting that future research should focus on the extent to which the pure Context-Driven view can be sustained or whether some interactionist version of the default view would be preferable.

In the next section we review the Gricean approach to conversational implicatures and we discuss its relation to the default and the Context-Driven approaches.

2. Grice and the context-driven approach

Grice (1989) posits two classes of inferences deriving from discourse: those based on the conventional meaning of the language used and those premised on the fact that a given utterance is made together with certain principles of discourse. The implications indicated in (2a,b) above represent typical cases of the latter type of inferences, normally referred to as 'conversational implicatures'. According to Grice, all such non-conventionally derived aspects of content could arise as nonce inferences premised on common expectations

concerning the speaker's behaviour — for example that s/he is being co-operative — and on more specific contextual assumptions.

As mentioned above, there are perceived differences between (2a) and (2b). Whereas (2b) seems strongly context-dependent, (2a) would seem to be available regardless of specific context. An important consequence of Grice's account of implicatures was that although an implication of an utterance might feel, intuitively, to be on a par with other more securely conventional components, it may not be conventional at all. This is the case with (2a).

The widely accepted semantic analysis of 'Some of the Fs have G' is that its meaning is consistent both with it being the case that some and not all of the Fs have G, but also with it being the case that all of the Fs have G. This can be illustrated by embedding the sentence in (1) within a conditional statement as in (3). Intuitively, (3) would be false if all of the students failed and the teacher kept her job:

3. If some of the students fail, then the teacher will be fired

But if the implication in (2a) were part of the conventional meaning of the sentence in (1), then (3) should be true under these circumstances.

To illustrate Grice's theory, we can consider briefly how the implication in (2a) above would be computed from first principles as an instance of Particularised Conversational Implicature. This nonce derivation stems from a conversational principle which enjoins the speaker to give as much information as is required for the current discourse purpose (this is usually referred to as a Quantity maxim). The derivation also turns on the presence in a given language of an alternative, equally economical form of words which would express a more informative statement. In the case of '*Some* of the Fs have G', '*All* of the Fs have G' is a minimally different more informative form. In cases where the extra information which the more informative forms encode would be relevant, the speaker would clearly be in breach of any Quantity principle by failing to use the manifestly more informative but equally economic form of words — unless s/he thought she could not give such information. If one can assume that the speaker would give the information if s/he could and that s/he is in a position to know whether or not all of the Fs have G, then one can infer that not all Fs have G.

Context-Driven accounts assume that where inferences of this variety are drawn, they are based on a nonce inference pattern of this type. Although Context-Driven theories may differ somewhat on the details of the derivation, they would assume that the interlocutors have to take into consideration (a) contextual assumptions as to what amount of information is required/relevant in a specific situation and (b) assumptions about each other's epistemic and volitional state.

The implicature in (2a) is often referred to as a *scalar implicature* due to the fact the alternative utterances 'Some/All of the Fs have G' form a scale of informativeness. Another commonly discussed scale is that involving 'or' and 'and'—illustrated in (4):

4. Mary dated John or Bill ⇒ Mary did not date both John and Bill

One can assume that 'or' has the conventional meaning of inclusive disjunction so that what (4) literally says is true even if Mary dated both John and Bill. The derivation of

the suggested implication proceeds as with 'some of the Fs', where the minimally different alternative involves the use of 'and' rather than 'or'.

Thus, according to the Context-Driven Account, these inferences are derived by taking contextual features into account. In this respect, there is no theoretically significant difference between Generalised and Particularised conversational implicatures. At best, the labels reflect superficial differences drawn from introspection. In the rest of the paper, we will generally refer to the inferences under study with the theory-neutral term, 'Scalar Implicatures' (SIs). Since part of what is being debated between the Default and the Context-Driven views is precisely whether 'Generalised Conversational Implicatures' is a term that describes a psychologically real type of inference, we will sometimes use the term 'GCI' when discussing the Default View.

3. Generalised conversational implicatures as default inferences

In contrast to the Context-Driven account, many proposals about common implicatures have suggested that these inferences are generalised in some particular sense. These include Gazdar (1979); Horn (1984); Levinson (2000), and Chierchia (2004).

According to this view, no nonce derivation of the type described in the last section occurs when the implicature triggers ('some of the Fs' or 'or') are encountered in the appropriate linguistic constructions[2]. Rather, these implicatures are considered prototypical cases of Generalised Conversational Implicature (GCIs) and are treated as a separate class of inference. The distinguishing feature of GCIs is their defaultness. That is, the inferences would be accepted unless there is directly conflicting evidence in the context.

> 'GCIs are inferences that appear to go through in the absence of information to the contrary; but additional information to the contrary may be quite sufficient to cause them to evaporate. Thus the mode of inference appears to have two properties: it is a *default* mode of reasoning, and it is *defeasible*.' Levinson (2000: 4, emphasis in original)

The conditions under which these default implicatures are thought to be implicitly cancelled include contexts such as that in (5) below (taken from Levinson, 2000: 51). In this context, A's question can be assumed to have explicitly fixed the level of relevance. It is enough for A to know that *at least* some of their documents are forgeries to get an answer to the question she raises. Whether or not all the documents are forgeries is irrelevant:

---

[2] Each proponent of the default view acknowledges that the inference in question is not triggered merely on encountering 'some of the F's' or 'or'. For example, each of these views would rule out the implicature in negative constructions. Exactly which linguistic constructions would or would not serve as a trigger is somewhat an open question - one which we leave to one side in comparing the Default and Context-Driven views. Suffice to say that our experimental studies utilise unequivocally triggering constructions according to any default view.

5 A: Is there any evidence against them?
  B: Some of their identity documents are forgeries.

According to Levinson, the scalar implicature, 'Not all of their documents are forgeries' is not available in this context (ibid). We will call such a context, where the default view supposes there is implicit cancellation a *lower-bound context* after the following quote due to Horn:

> 'a generalised implicatum … may be cancelled-explicitly (cf. *Some, {if not all, and perhaps all,} of my friends…*) or implicitly (by establishing the appropriate contexts, in which all that is relevant, or can be known, is the lower-bound).' (Horn, 1984: 13)

Among the researchers who share the Default View, Chierchia (2004) has proposed that GCIs are derived according to the compositional semantic rules of the grammar and Levinson (2000) has posited a specialised pragmatic system which yields generalised pragmatic inferences.

4. Experimental studies comparing default and context-driven accounts

On the question of how scalar implicatures are processed on-line, proponents of both the Context-Driven and Default views have emphasised the importance of experimental validation. Levinson and Chierchia are explicit in claiming that generalised scalar implicatures are expected to be *generated* automatically once a triggering construction has been processed and before full sentence meaning is computed. Since scalar implicatures are defeasible, context may suggest at a later stage that a certain inference was incorrect and cancel the implicature (perhaps through a process similar to retracting an analysis of a garden path sentence). On the Context-Driven side, Sperber and Wilson's and Carston's account of pragmatic inference implies that although scalar implicatures are generated on-line, this only happens when the cognitive representation of context warrants it (Carston, 2004; Sperber and Wilson, 1995; Wilson and Sperber, 2003). Moreover, Sperber and Wilson (2002) have suggested that pragmatic implicature is a product of attentional processing, requiring effort beyond that devoted to automatic, default processes.

Given the above discussion of scalar implicatures, there are two sets of differing predictions made by these competing accounts. One concerns the amount of resources required to process the implicature triggers and the other concerns the availability of the implications themselves.

The predictions are concerned with three kinds of context. In the last section, we referred to the context in (5) as a *lower-bound context* since it is clear that only the lower-bound interpretation (i.e. without the implicature) is relevant. An *upper-bound context* is one where it is clear that the upper-bound interpretation (i.e. with the implicature) is relevant. For example, an upper-bound context for B's utterance in (5) would contain a question like 'Were all their documents in order?'. It is also possible to conceive of B's utterance occurring in a context where neither the question of whether some documents were in order nor the question of whether all the documents were in order is at issue. We will call such contexts *neutral contexts*.

According to both Context-Driven and Default approaches, no scalar implicature would be available in lower-bound contexts. Both positions also agree that a scalar implicature is available in an upper-bound context. So on the question of availability, in both lower-bound and upper-bound contexts both the Default and Context-Driven accounts predict the same outcome. The real difference between the two positions comes with neutral contexts. The Default view claims that generalised scalar implicatures are generated by default and hence they are generated in neutral contexts. The Context-Driven view claims that scalar implicatures need contextual support. No such support is available in neutral contexts and hence no implicature will be available or generated in neutral contexts. We test these predictions in the second set of experiments reported below.

Returning to upper-bound and lower-bound contexts, even though both approaches predict the same outcome with regard to whether or not an implicature is available, the critical difference between the two positions concerns *how* implicatures are generated (and cancelled). According to the Default view, scalar implicatures are triggered by linguistic constructions themselves in any contexts and are generated without any difference in inferential resources. However, these inferences need to be cancelled somehow in lower-bound contexts. By contrast, the Context-Driven view suggests simply that no inference is generated in lower-bound contexts. In upper-bound contexts, the Context-Driven account suggests that the implicature is generated via a process which demands that extra attention be devoted both to the level of information which is relevant in the context and also to the speaker's epistemic state.

For the case of 'or', (6a,b) illustrate upper-bound and lower-bound contexts for the same triggering phrase, 'the class notes or the summary':

6a. (*Upper-bound*) John was taking a university course and working at the same time. For the exams he had to study from short and comprehensive sources. Depending on the course, he decided to read *the class notes or the summary*.
 b. (*Lower-bound*) John heard that the textbook for Geophysics was very advanced. Nobody understood it properly. He heard that if he wanted to pass the course he should read *the class notes or the summary*.

Given texts like (6a,b) which include identical triggering expressions, we should expect there to be some difference in the predictions for reading times of the triggering expressions between lower- and upper-bound contexts. We can safely assume that any additional inferential activity on a word or expression is reflected in additional processing time (e.g. O'Brien, Duffy, & Myers 1986). The Context-Driven approach predicts that experimental subjects would tend to derive the scalar implicature in the upper-bound context while not doing so in the lower-bound contexts, and hence predicts a positive difference between the reading times for the trigger phrases in (6a,b).

There are two possible predictions for the default model. If cancellation of a default inference is a process that requires extra processing time, the default view predicts a longer reading time in lower-bound contexts and hence a negative difference between contexts. If cancellation is merely a matter of decay of an activated inference, then it is expected that there will be no difference between upper- and lower-bound contexts. Either way, the two

models make distinct predictions about the differences in reading time for the triggers (positive vs. negative/no difference).

Previous studies have compared response times between cases where scalar implicatures are drawn and when they are not. Noveck and Posada (2003); Bott and Noveck (2004) asked subjects to perform a truth-value judgement task on sentences like (7):

7. Some elephants have trunks

Where subjects judge the sentence false, they are basing that judgement on the upper-bounded interpretation, 'Some and not all elephants have trunks'. That is, the interpretation includes the scalar implicature. Where the response is 'true', this would be based solely on an interpretation drawn from the linguistic meaning.

Noveck and Posada (2003) timed responses and found that the time taken by the participants answering 'false' was significantly longer than those answering 'true'. This suggests that when respondents based their answers on the plain meaning, it was not the case that the scalar implicature was first generated and then cancelled but rather that the implicature was not generated in the first place. This result conforms to the predictions of the Context-Driven view.

Bott and Noveck (2004) replicated these findings in an experiment where an additional layer of narration is introduced. The stimuli were preceded either by the declaration that "Mary says the following sentence is true" or that "Mary says the following sentence is false". In this way, one can elicit the same kind of response to each type of interpretation (with or without the implicature). In another experiment Bott and Noveck found that the number of responses based on the scalar implicature increased as permitted response time increased. All of these results suggest that scalar implicatures are costly inferences in terms of how much time they need in order to be generated and that they manifest none of the automaticity or speed that the Default view would expect. Therefore the studies by Noveck and colleagues have yielded consistent evidence against the existence of default implicatures. As they note, the results favour the Context-Driven view which assumes that scalar implicature generation is a fully fledged inferential process that comes with a processing cost.

By contrast, Bezuidenhout and Cutting, 2002, henceforth B&C) report that in a text comprehension experiment where triggers occur in upper- and lower-bound contexts, the time taken to read the triggers was longer in lower-bound contexts. This result supports the Default view, conflicting with the results obtained by Noveck and colleagues. However, B&C themselves raise questions about the items they used (ibid: 452). They drew many of the materials for their experiment from items used in other off-line studies (those of Gibbs & Moise, 1997; Nicolle & Clark, 1999) which were not designed to test scalar implicature. As a result, nearly half of the triggers used were not unequivocally implicature triggers.

The problematic set of triggers that B&C use include cardinal phrases such as 'two children'. These have traditionally been thought to mean simply 'at least two children'. Where understood to mean 'exactly two' as in (8), this was thought to be due to the presence of a scalar implicature to the effect that Mary has at most two children:

8. Mary has two children.

However, there has never been complete agreement about whether the 'exactly' understanding of cardinals is really due to an implicature. They are not considered implicatures by relevance theorists (see Carston, 1998; Carston, 2002; Papafragou & Musolino, 2003 for developmental evidence). Even Levinson (2000: 90) notes that 'the cardinals are *not* the correct test bed for the whole theory of…implicature' (italics in original). Horn (1992); Geurts (1998a) hold a similar view. On each of these dissenting views, the 'exactly' and 'at least' interpretations of cardinals are on a par. If anything, the 'exactly' interpretation is the dominant one.

B&C's items also included possessive phrases such as 'John's book' and quantified noun phrases such as 'every student'. B&C assumed that inferring the relation between John and the book or the domain of quantification could be counted as deriving an implicature. However, there is little theoretical support in favour of this assumption. Levinson (2000) seems to be the only theorist claiming that possessives involve an instance of implicature and not even he would claim that setting the domain of a quantifier involves an implicature. The current consensus is that possessive phrases and quantification are arguably different from implicature triggers since, although they involve pragmatic inference for their interpretation, the inference is mandated by the linguistic form itself (see Geurts, 1998b; Stanley, 2000; Westerstahl, 1985). That is, while the interpretation of 'every student' involves determining what set of students is being quantified over, it is agreed that these expressions contain some kind of hidden parameter in the logical form which has to be fixed in the context. Given that possessives and quantified noun phrases contain a hidden parameter, then the same kind of pragmatic inference is always mandated for their interpretation. However, B&C assumed that there was a lack of implicature when 'every student' was understood to mean, 'every student in the world'. Arguably, the contexts used to elicit such interpretations are more marked, and hence more difficult to access than those where the implicit domain variable is assigned to a set made salient in the context. So, there is reason to call B&C's results into question on the basis of the materials used.

In short, the picture that is drawn from the small, but growing literature on scalar implicature on-line processing is not conclusive: single sentence truth value judgment tasks by Noveck and colleagues consistently support the Context-Driven approach, but text-comprehension studies by Bezuidenhout and colleagues are equivocal.[3]

5. Experiment 1: replicating Bezuidenhout and Cutting

In our first experiment, we performed the same kind of study as Bezuidenhout and Cutting's except we used triggers which uncontroversially involve implicature. We measured the reading times of a trigger-containing segment at the end of a context which

---

[3] Bezuidenhout and Morris (2004) report results of a study which also point in favour of the Context-Driven view. However, their conclusions are based on a notion of underspecification which we have argued is problematic (see Katsos, Breheny and Williams 2005).

Table 1
The English translation of an item from Experiment 1

Upper-bound context
John was taking a university course/and working at the same time./For the exams/he had to study/from short and comprehensive sources./Depending on the course,/he decided to read/the class notes or the summary./
Lower-bound context
John heard that/the textbook for Geophysics/was very advanced./Nobody understood it properly./He heard that/if he wanted to pass the course/he should read/the class notes or the summary./

Note: The critical phrase is in italics, but was presented in the standard font in the experiment.

biases the scalar implicature and at the end of a context, where the implicature is intuitively not available because 'all that is relevant is the lower-bound'. Instead of testing different categories of items we only tested for one trigger, namely the disjunction 'or'. As stated above, the semantic interpretation of the disjunction 'A or B' is the inclusive reading and the implicature is the exclusive 'either A or B but not both'.

Participants were given discourses like those in (6a,b), repeated in Table 1, in a segment-by-segment self-paced reading task. The experiment was conducted in Greek.[4] Slashes indicate the segmentation of the text.

5.1. Method

5.1.1. Participants

The participants were 47 native speakers of Greek (aged between 19 and 25 years; mean age 23.5 years). There were 25 males and 22 females. They were students at the University of Cambridge, studying a range of science and arts subjects (but none was taking linguistics courses). All participants had normal or corrected-to-normal vision. They were paid 3 British pounds for their participation.

5.1.2. Materials

A total of 12 pairs of texts like those in Table 1, but written in Greek, comprised the critical materials. The texts were between 32 and 40 words in length; texts of the same pair did not differ in size by more than four words. The trigger phrase always occurred in the last sentence, at the end of the sentence. The 12 pairs of texts were randomly divided into two sets of 6. One presentation list was created by taking the Upper-Bound texts from one set and the Lower-Bound texts from the other set. A second presentation list was created using the remaining versions of each text. Each presentation list, therefore, contained six items in each condition. Each presentation list was also supplemented with 42 filler texts which did not contain the inference under study, but were the critical items for a different experiment. The filler texts were similar in size to the critical items. The order of the texts in each list was independently randomised. In order to motivate comprehension, 33% of all the texts (both critical and fillers) were followed by a yes/no comprehension question. The answer to the question was contained in the text that immediately preceded it.

---

[4] The Greek disjunctive particle 'i' has similar properties to its English counterpart 'or' (it has an exclusive and inclusive interpretation).

An affirmative answer required pressing the left mouse button and a negative pressing the right mouse button. Half of the answers required an affirmative answer and half a negative one. The questions did not concern the target inference and came randomly during the presentation sequence (see Appendix for sample items and a comprehension question). Materials for three practice items which did not contain the target inference were also presented. One of the practice items had a question.

5.1.3. Procedure

The experiment was run on a laptop with an LCD display using Superlab software. Participants were told that they had to answer accurately some questions pertaining to sentences that they would read on the computer. They were told that they would read short texts presented segment by segment. After reading each segment carefully they were told to press the left mouse button to see the next one using their habitual mouse hand. They were instructed to read each segment only once and not try to memorise it, but also that they should fully understand each segment before advancing to the next one. The segments were presented in black font against a white screen in a fixed point near the centre of the screen. There was no delay between the pressing of buttons and the presentation of segments. The time to read each segment was recorded by the software. Answers to comprehension questions were recorded by the software as well. After the end of each sentence and before the next sentence began we introduced a 1000 ms interval to allow participants to complete all inferential activities before starting to read the next sentence. The experiment took between 20 and 25 min to complete, including one rest break half way through. The post-experiment interview revealed that no participant had become aware of the purpose of the experiment. Participants' accuracy to the comprehension questions was 92%, with accuracy rates being similar in questions following the critical and filler items. However, three participants who scored lower than 80% were excluded from the final sample and replaced.

5.2. Results and discussion

The reading times for the trigger segments constituted the critical data for analysis. The effect of outliers was curtailed in two stages. First, excessively long reading times, assumed to be due to processes other than the ones of interest, were excluded. The exclusion value was calculated as approximately 2.5 times the mean critical segment time over all participants. In the present case, this was 3000 ms. Also, any values below 200 ms were excluded, since they probably reflected mechanical responses. Once this had been done, individual cut-off values were calculated for each participant as the mean plus 2 standard deviations over all 12 critical items; that is, disregarding condition. Any value exceeding the cut-off was 'smoothed' to the cut-off value. The assumption was that such long responses may reflect the processes of interest, but that their effects should be curtailed in order that the response time distribution not be overly skewed. Over all participants, 2.48% of the data were either removed or replaced by the cut-off value for each condition.

The mean trigger segment reading times for the Upper-Bound and Lower-Bound contexts were 1291 (SD=352) and 1204 (SD=292) ms, respectively. Analyses of variance were conducted in which subjects and items were treated as random factors.[5] These analyses revealed a significant main effect of Condition, $F1(1,43)=7.65$, $P<0.01$; $F2(1,10)=6.61$, $P<0.05$. Reading time on the implicature trigger in Upper-Bound contexts was significantly longer than in the Lower-Bound contexts.

As predicted by Context-Driven approaches, reading time was significantly longer in the condition where the implicature is warranted by the context, indicating that implicatures are generated only in such cases. If implicatures were generated in contexts that did not warrant them and were subsequently cancelled, this should have manifested as longer reading time on the lower-bound context condition or at least as equal reading time between conditions. This study shares some of the original B&C study's limitations: for example, the triggers are preceded by different segments in either context and this may have an effect on the reading time on the following segment (the trigger) that neither we nor B&C controlled for. Note however, that even though we used a similar methodology to B&C, we obtained the opposite results. We believe that this difference is due to the fact that we used a single type of uncontroversial scalar implicature trigger.

6. Experiment 2: scalar implicature generation in neutral contexts

The second set of differing predictions made by the Default and Context-Driven views concern what happens in genuinely neutral contexts. A neutral context for scalar triggers is just a context, where there is no information about whether all that is relevant is the lower-bound and no information about whether all that is relevant is the upper-bound. Our second study attempts to probe the generation of 'some of the Fs G-ed' > 'not all the Fs G-ed' scalar implicatures (cf (1b) above) in neutral contexts.

One might think that the most obvious kind of example to illustrate neutral contexts would be an out-of-the-blue utterance containing a scalar implicature trigger. However, as is well known, even single sentence utterances can create their own context through a variety of presupposition triggers and information-structure triggers. For example, 'again' is a presupposition trigger: even if you did not know that John had been to Tibet, an utterance of 'John is going to Tibet again' would result in the accommodation of such contextual information.

An example of an information-structure trigger would be a clefted sentence as exemplified in (10).

10. It was John who ate the cake

In this example, the noun phrase 'John' is linguistically marked as focus — the answer to an implied contextual issue: *Who ate the cake?*. Again, if this issue is not

---

[5] For the subjects analysis the between-subjects factor was Presentation List, and the within-subjects factor was Condition (Lower-Bound, Upper-Bound). For the items analysis the between-items factor was Item Set and the within-items factor was Condition.

already salient for the hearer, an utterance of (10) would trigger its accommodation into the context.

It is possible to manipulate information-structural properties of sentences containing 'some of the Fs' so that out-of-the-blue tokens would be quite likely to trigger the accommodation of a contextual issue such as *What happens with the Fs?*. It is also possible to construct examples where this accommodation is less likely. Where a question like, *What happens with the Fs?* is accommodated into a context, that context would be an upper-bound context for the trigger 'Some of the Fs'. An upper-bound context is one where the scalar implicature is available according to both accounts. Where no such issue is accommodated for an out-of-the-blue case, this would constitute a genuinely neutral context for the trigger. According to the default view, the scalar implicature is available in neutral contexts; according to the Context-Driven view, it is not. In Experiment 2, we constructed materials with the aim of manipulating the salience of the relevant upper-bound issue.

To explain our design, let us adopt a simple model of utterance contexts containing facts and issues (see van Kuppevelt, 1996; van Rooy, 2004). The normal expectation is that utterances will resolve issues in such contexts. So-called 'old information' in sentences relates to contextual issues. For example, if an issue is *Who did Mary kiss?* then in the sentence 'Mary kissed Bill' the frame 'Mary kissed_' is old information. By contrast, if an issue is *Who kissed Bill?* then '_ kissed Bill' is old information. English speakers often mark new information with focal stress. As in many other languages, English speakers tend to construct sentences so that old information comes at the beginning of the sentence. Thus, out of the blue, a sentence is most likely to be understood as addressing a contextual issue determined by the earlier parts of the sentence. So 'Mary kissed Bill' will more likely be understood to be about what happened with Mary or about who Mary kissed rather than being about who kissed Bill. Accordingly, an out-of-the-blue token of (11a) is more likely to be related to an issue like *What happened with the consultants?* or *Who did the consultants meet with?* than is (11b):

11a. Some of the consultants had a meeting with the director
  b. The director had a meeting with some of the consultants

The idea for our design was to construct pairs such as those in (11), where there is a difference in the likelihood that the trigger 'some of the Fs' is understood by experimental participants as relating to a contextual issue which involves the Fs in this way.

For reasons having to do with the fact that English is not a flexible word-order language, this old-information-first tendency is quite mild in English. However, things are far more congenial in Greek, since Greek has flexible word order. That is, both grammatical subjects and objects can appear first in an utterance. This gives rise to a far stronger tendency in Greek for grammatical subjects in sentence-initial position to be construed as old information and objects in a non-initial, post-verbal position to be construed as unrelated to a contextual issue (see Philippaki-Warburton, 1987 i.a.).

For our study, we created short two-sentence discourses in Greek (like those in (12a,b)). In the (a) example, the trigger 'some of the Fs' is a subject in pre-verbal position (i.e. in an

SVO order). In the (b) example, the trigger 'some of the Fs' is the grammatical object in post-verbal position.

12a. Meriki apo tus simvulus ihan sinadisi me to diefthidi.
      Some-nom pl. of the consultants-acc. had-3rd pl. past meeting-acc. with the director-acc.
      I ipolipi den kataferan na parevrethoun.
      The rest-nom pl. neg managed-3rd pl. attend-3rd pl. past. subj
      (Some of the consultants had a meeting with the director. The rest did not manage to attend.)
  b. O diefthidis ihe sinadisi me merikus apo tus simvulus.
      The director-nom. had-3rd pl. meeting-acc. with some-acc pl. of the consultants-acc.
      I ipolipi den kataferan na parevrethoun.
      The rest-nom pl. neg managed-3rd pl. attend-3rd pl. past. subj
      (The director had a meeting with some of the consultants. The rest did not manage to attend.)

In both (12a) and (12b), there is a second sentence which always started with the target phrase 'the rest' (or 'the others'). The target phrase's interpretation can be recovered immediately but requires making a bridging inference that the collection referred to by the 'the rest' is the rest of the Fs. What is important is that the bridging inference will be facilitated if the inference 'not all of the Fs' has been generated on-line in the first sentence. When the implicature has not been generated, we expect longer reading time on the target for two reasons: first, because the anaphoric expression's interpretation has to be fully computed on the spot without any facilitation from previous inferencing; and second, because some re-analysis may take place on the trigger (to the effect that 'some' is reinterpreted to implicate 'not all').

Given our assumptions about the effect of the position of the triggering phrase on any accommodated issues, the Context-Driven view predicts that sentence position should have an effect on the reading time for 'the rest'. This is because, if there is no accommodation of an appropriate contextual issue, there is no reason to draw the scalar implicature in the neutral context. Thus reading time should be facilitated by the trigger being in sentence-initial position. By contrast, default processing models predict that regardless of where the implicature trigger is, the implicature should be generated on-line. So, on the default view, sentence position should not affect reading time.

Participants were given discourses like those in Table 2 (presented here in English for ease of exposition. Slashes indicate segmentation). As Table 2 indicates, we added a second *Explicitness* factor, '*some*' vs '*only some*' in addition to the *Position* factor, *Sentence initial* vs *Sentence final*.

In both of the '*only some*' conditions the 'not all'-inference is made explicit by the presence of the operator 'only'. This operator explicitly encodes the inference and should make reading time on the phrase 'the rest' as fast as possible. The '*only some*' conditions were necessary as controls since (a) the distance between the trigger 'some of the Fs' and the target phrase 'the rest' is unequal in the two '*some*' conditions; and (b) the trigger in

Table 2
The English translation of an item from Experiment 2

Sentence initial-'some'
Some of the consultants/had a meeting/with the director./The rest/did not manage/to attend./
Sentence initial-'only some'
Only some of the consultants/had a meeting/with the director./The rest/did not manage/to attend./
Sentence final-'some'
The director/had a meeting/with some of the consultants./The rest/did not manage/to attend./
Sentence final-'only some'
The director/had a meeting/with only some of the consultants./The rest/did not manage/to attend./

Note: The target phrase is in italics (but was presented in the standard font in the experiment).

the *Sentence initial* condition agrees in case with the target (nominative-nominative), but there is no agreement between trigger and target in the *Sentence final* condition (accusative-nominative). Potential effects from these two factors are neutralised by comparing pairs of '*some*'-'*only some*' conditions.

In both '*some*' conditions, the 'not all'—inference would rest solely upon the scalar implicature. Reading time for 'the rest' should be facilitated if the 'not all'—inference is made on-line. So we can assume that an implicature is generated on-line if reading time on the target phrase 'the rest' is comparable in the '*some*' and '*only some*' conditions.

6.1. Method

6.1.1. Participants

The participants were 48 native speakers of Greek, (aged between 20 and 24 years; mean age 22 years). There were 22 males and 26 females. They were students at the University of Athens, studying a range of science and arts subjects (but none was taking linguistics courses). All participants had normal or corrected-to-normal vision. They were paid 3 Euros for their participation.

6.1.2. Materials

A total of 24 sets of texts, each in four conditions, like those in Table 2 but written in Greek, comprised the critical materials. The texts were between 15 and 20 words in length; texts of the same set did not differ in size by more than 4 words. The trigger phrase 'some of the Fs' always occurred in the first sentence. In the *Sentence initial*-'*some*' and -'*only some*' conditions the trigger appeared as the first segment of the sentence; in the *Sentence final*-'*some*' and -'*only some*' conditions it appeared as the last segment of the sentence. The target segment was '/the rest/' in half of the items and '/the others/' in the other half of the items. The target always appeared in the first segment of the second sentence. The target segment was followed by a segment that was identical in all four conditions, so that potential spill-over or delayed effects could be measured. Note that in the first sentence there were no grammatical or conceptual plurals besides the phrase 'some of the Fs', because their presence might have even slightly interfered with the reference assignment process. The critical items were created so that the only possible referent of the target phrase 'the rest/the others' was always 'the Fs which are not G'. We also took care to

eliminate materials which might independently trigger scalar implicatures. These include the use of predicates such as 'pass'/'fail' which would prompt readers to presuppose that the question of whether all passed/failed is at issue.

The 24 sets of texts were randomly divided into four groups of 6. One presentation list was created by taking the *Sentence initial-'some'* texts from Group 1, the *Sentence initial-'only some'* texts from Group 2, the *Sentence final-'some'* texts from Group 3 and the *Sentence final-'only some'* texts from Group 4. Another three presentation lists were created using the remaining versions of each text in a similar manner. Each presentation list, therefore, contained six items in each condition. Each presentation list was also supplemented with 85 filler texts which did not contain the inference under study. The filler texts were similar in size to the critical items. The order of the texts in each list was independently randomised.

In order to motivate comprehension, 25% of all the texts (both critical and fillers) were followed by a yes/no comprehension question. The answer to the question was contained in the text that immediately preceded it. An affirmative answer required pressing the left mouse button and a negative answer the right mouse button. Half of the questions required an affirmative answer and half a negative one. The questions did not concern the target inference and came randomly during the presentation sequence (see Appendix for sample items and a comprehension question). Materials for three practice items which did not contain the target inference were also presented. One of the practice items had a question.

6.1.3. Procedure

The procedure was identical to that of Experiment 1. The experiment was run on a laptop with an LCD display using Superlab software. Participants were told that they had to accurately answer some questions pertaining to the sentences that they would read on the computer. They were told that they would read short texts presented segment by segment. After reading each segment carefully they were told to press the left mouse button to see the next one using their habitual mouse hand. They were instructed to read each segment only once and not try to memorise it, but also that they should fully understand each segment before advancing to the next one. The segments were presented in black font against a white screen at a fixed point near the centre of the screen. There was no delay between the pressing of buttons and the presentation of segments. The time to read each segment was recorded by the software. Answers to comprehension questions were recorded by the software as well. After the end of each sentence and before the next sentence began we introduced a 1000 ms interval to allow participants to complete all inferential activities before starting to read the next sentence. The experiment lasted approximately 30 min, including a rest break half way through. The post-experiment interview revealed that three participants had become aware of the pattern of the critical items. An analysis that was performed excluding these participants showed that the significance of the results remains the same with or without them. Participants' accuracy on the comprehension questions was 93%, with accuracy rates being similar in questions following the critical and filler items. However, two participants who scored lower than 80% on the questions were excluded from the final sample and replaced.

6.2. Results and discussion

The raw data were treated in the same way as in Experiment 1. The critical value for exclusion was set at 1500 ms. Over all participants, 3.47% of the data were removed and smoothed for the *Sentence initial-'some'* condition, 2.43% for the *Sentence initial-'only some'* condition, 3.47% for the *Sentence final-'some'* condition, and 2.77% for the *Sentence final-'only some'* condition. The mean target segment reading times in these conditions are shown in Table 3.

Analyses of variance were conducted in which subjects and items were treated as random factors. These analyses revealed no significant effect of *Sentence Position*, $F1$ and $F2$ both <1.0. There was a main effect of *Explicitness* in the subjects analysis, $F1(1,44)=4.94$, $P<0.05$, but not in the items analysis, $F2(1,19)=2.64$, $P>0.1$. The critical interaction between *Sentence Position* and *Explicitness* was significant, $F1(1,44)=4.24$, $P<0.05$; $F2(1,19)=6.93$, $P<0.05$. As can be seen from the means in Table 3, in the *Sentence final* condition reading times were slower in the '*some*' case than the '*only some*' case, but there was no such difference in the *Sentence initial* conditions. Also, there was no reliable difference between any other comparable segments of the texts, including the segments following the target phrase.

The results show a significant *Sentence initial* vs *Sentence final* asymmetry. This suggests that implicatures are drawn on-line for *Sentence initial-'some'* triggers but not for *Sentence final-'some'* triggers.

We suggest that these results conform to the predictions of the Context-Driven view. However, there are two points arising from these results that call for some discussion. First, the fastest reading time was obtained on the *Sentence final-'only some'* condition rather than on any of the *Sentence initial* conditions. Although this might seem unexpected at first sight, it is explained if we consider the distance between the scalar implicature trigger and the target phrase "the rest" in each condition. In the *Sentence initial* conditions, when the target phrase is processed it must be linked to information that is contained in the trigger which is separated from the target by two intervening segments. A certain slowdown can be expected due to the target-trigger distance measured in number of words or entities introduced. By contrast, in the *Sentence final* conditions the trigger segment is adjacent to the target phrase and no other entity is introduced between them. Therefore distance penalties are not expected.

Second, a further explanation is required for the slowest reading time that was obtained in the *Sentence final-'some'* condition. A potential objection to how we interpret these results could be that the slow reading time is due to a penalty of dispreferred continuation

Table 3
Mean reading time on the target segment 'the rest' or 'the other(s)' in Experiment 2

	Reading time (milliseconds)	Standard deviation
Sentence initial-'some'	613	125
Sentence initial-'only some'	611	110
Sentence final-'some'	628	138
Sentence final-'only some'	586	112

between the first and second sentence rather than scalar implicature generation. The objection would go like this: When participants read the first sentence of each item they might assume that the preferred continuation for the forthcoming second sentence is the first NP mentioned in the first sentence. In both the *Sentence initial* conditions the target phrase "the rest" in the second sentence makes reference to part of a set introduced by the first NP mentioned in the first sentence; so in both conditions the sentences are linked with reference to a preferred continuation and therefore both conditions are comparable. In both the *Sentence final* conditions, the target phrase "the rest" refers to a set introduced by the *second* NP of the first sentence, thus violating preferred continuation expectations. Again, both *Sentence final* conditions are comparable, and both could be paying a potential penalty. However, it is possible that in the *Sentence final-'only some'* condition, the presence of "only" may rearrange the order of preferred continuations. It may well be that an NP modified with the focus operator "only" will become the preferred point of continuation even if it is the second NP to be mentioned in a sentence. In this case, the *Sentence final-'only some'* condition is not violating preferred continuation expectations. Under these assumptions, there is only one condition that is violating preferred continuation expectations, the *Sentence final-'some'* condition, and thus the reading time slowdown that was obtained on this condition may have to be attributed to such a penalty.

To investigate whether this objection is valid, we ran an off-line sentence continuation task using exactly the items of Experiment 2 in the *Sentence final-'some'* and the *Sentence final-'only some'* conditions. Twentyone native speakers of Greek were presented with just the first sentence of the 24 items and were asked to fill in one more sentence that would be the continuation. Another 20 unrelated filler sentences were added. In the *Sentence final-'some'* condition 30% of the continuations referred to sets introduced by the first sentence's initial noun phrase, NP1; and 52.5% of continuations referred to sets introduced by its final noun phrase, NP2. In the *Sentence final-'only some'* condition the proportion of continuations referring to sets introduced by the first sentence's NP1 was 24.2% and by its NP2 was 59.2%. Crucially, when comparing NP1 and NP2 responses, the interaction between *Explicitness* ('some', 'only some') and type of *Continuation* (NP1, NP2) was not significant ($F = 2.41$) but there was a main effect of *Continuation*, $F(1, 17) = 7.4$, $P < 0.05$. An analysis of variance on simple main effects revealed that in both conditions it was the final NP continuation that was preferred over the initial NP (for '*some*' $F(1, 16) = 5.01$, $P < 0.05$; for '*only some*' $F(1, 16) = 12.09$, $P < 0.01$).

These results suggest that indeed sentence final NPs in the *Sentence final-'only some'* condition were rearranged at the top of the order of preferred continuations but so were sentence final NPs in the *Sentence final-'some'* condition. So in neither *Sentence final* condition was there any dispreferred continuation penalty. The difference that was obtained between the two conditions can be attributed to the fact that in the *Sentence final-'only some'* condition, due to the presence of 'only', both the set of the Fs that G-ed and the set of the Fs that did not G were made available; whereas in the *Sentence final-'some'* condition only the set of the Fs that G-ed was available and therefore a slowdown was required when processing the anaphoric expression "the rest of the Fs" that refers to the set of the Fs that did not G.

7. Experiment 3: controlling for plugs

At this point, it could perhaps be argued by proponents of default theories regarding Experiment 2 that Greek sentence structure is one of the determinants in the (yet to be specified) system of GCI 'plugs' — i.e. linguistic constructions which specifically block the default inference. That is, it could be that implicatures were not generated in sentence-final positions in Greek because in that position they are always blocked. However, this seems unlikely. According to most linguistic accounts, the object position in an SVO sentence in Greek is distinguished only by an absence of information-structural features (see Phillipaki-Warburton, 1987). One would imagine that rules for GCI plugs involve the presence of some feature (such as negation). Nevertheless, in our third experiment, we set out to investigate if it is indeed the case that scalar implicatures are blocked in sentence-final positions. To do so we manipulated contextual expectations. We put the trigger 'some of the Fs' in sentence-final position and we embedded it in discourses which were either upper-bound contexts or lower-bound contexts. We also included a baseline condition with the operator 'only', as in Experiment 2. We ran the experiment in Greek as well. Items are represented in Table 4 with further examples in the appendix (slashes indicate segmentation).

Two reading time measures can be taken in this design, one on the trigger phrase 'some of the Fs', similar to Experiment 1 and one on the target phrase 'the rest', similar to Experiment 2[6]. Given our previous findings, when the scalar implicature has been generated we expect longer reading time on the trigger—an inference requires additional time to be generated - but shorter reading time on the target - the inference has already been made so the anaphoric phrase's interpretation is facilitated. On the other hand, when the implicature has not been generated we expect shorter reading time on the trigger and longer reading time on the target. We assume that the *Upper-Bound-with-'only some'* condition (where the scalar implicature is facilitated by context and is also explicitly encoded by the operator 'only') is our baseline for judging if the implicature has been generated in any of the -*with* '*some*' conditions.

The Context-Driven approach predicts that the scalar implicature will be generated in the '*some*' and '*only some*' upper-bound context condition, but it will not be generated in the '*some*' lower-bound context condition. A Revised Default model, with blocks on implicatures in sentence-final positions, would predict that the scalar implicature is not generated in either of the '*some*' conditions (with upper- or lower-bound contexts). The standard Default model we discussed in Experiments 1 and 2 predicts that the implicature is generated in all three conditions but cancelled in the lower-bound condition.

---

[6] Note that in Experiment 2 the trigger segments are not included in the critical data because they are not comparable (they are in different positions in the sentence, initial vs. final, and they are explicitly marked for different case in Greek, nominative vs. accusative).

Table 4
The English translation of an item from Experiment 3

Upper-Bound Context with 'some'
Mary/asked John/whether he intended to host/all his relatives/in his tiny apartment./John replied/that he intended to host/some of his relatives./The rest/would stay/in a nearby hotel.
Upper-Bound Context with 'only some'/
Mary asked John/whether he intended to host/all his relatives/in his tiny apartment./John replied/that he intended to host/only some of his relatives./The rest/would stay/in a nearby hotel.
Lower-Bound Context with 'some'/
Mary was surprised/to see John/cleaning his apartment/and she asked/the reason why./John told her/that he intended to host/some of his relatives./The rest/would stay/in a nearby hotel.

Note: The critical phrases are in italics (but were presented in the standard font in the experiment).

7.1. Method

7.1.1. Participants

The participants were 31 native speakers of Greek, aged between 20 and 25 years; mean age 23.5 years. There were 16 males and 15 females. They were students at the University of Athens, studying a range of science and arts subjects (but none was taking linguistics courses). All participants had normal or corrected-to-normal vision. They were paid 3 Euros for their participation.

7.1.2. Materials

A total of 18 sets of texts, each in three conditions, like those in Table 5, but written in Greek, comprised the critical materials. The texts were between 33 and 41 words in length; texts of the same set did not differ in size by more than 4 words. The trigger phrase 'some of the Fs' always occurred as the final segment of the first sentence. The target segment was the phrase 'the rest' in half the items and 'the others' in the other half. The target segment appeared as the first segment of the second sentence. Target segments were followed by a segment that was identical in all three conditions, so that we could track whether slower reading time on the trigger in any condition would carry on to the following segment as well. Note again that in the critical texts, in the first sentence there were no grammatical or conceptual plurals besides the phrase 'some of the Fs' and that we did not include materials which might independently trigger scalar implicatures.

The 18 sets of texts were randomly divided into three groups of 6. One presentation list was created by taking the *Upper-Bound with 'some'* texts from Group 1, the *Upper-Bound with 'only some'* texts from Group 2 and the *Lower-Bound with 'some'* from Group 3. Another two presentation lists were created using the remaining versions of each text in

Table 5
Mean segment times in milliseconds and standard deviations (in parentheses) for the triggers and targets in Experiment 3

	Upper- Bound with 'some'	Upper- Bound with 'only some'	Lower- Bound with 'some'
Trigger segment	1027 (242)	-	927 (184)
Target segment	604 (114)	612 (144)	655 (168)

a similar manner. Each presentation list therefore contained 6 items in each condition. Each presentation list was also supplemented with 36 filler texts which did not contain the inference under study but were the critical items of a different experiment. The filler texts were similar in size with the critical items. The order of the texts in each list was independently randomised.

In order to motivate comprehension, 25% of all the texts (both critical and fillers) were followed by a yes/no comprehension question as in previous experiments (see Appendix for sample items and a comprehension question). Materials for three practice items which did not contain the target inference were also presented. One of the practice items had a question.

7.1.3. Procedure

The procedure for this experiment is identical to Experiments 1 and 2. The experiment was run on a laptop with an LCD display using Superlab software. Participants were told that they had to accurately answer some questions pertaining to sentences that they would read on the computer. They were told that they would read short texts presented segment by segment. After reading each segment carefully they were told to press the left mouse button to see the next one using their habitual mouse hand. They were instructed to read each segment only once and not try to memorise it, but also that they should fully understand each segment before advancing to the next one. The segments were presented in black font against a white screen at a fixed point near the centre of the screen. There was no delay between the pressing of buttons and the presentation of segments. The time taken to read each segment was recorded by the software. Answers to comprehension questions were recorded by the software as well. After the end of each sentence and before the next sentence began we introduced a 1000 ms interval to allow participants to complete all inferential activities before starting reading the next sentence. The experiment took approximately 20 min to complete, including one rest break half way through. The post-experiment interview revealed that two participants had become aware of the pattern of the critical items. An analysis that was performed excluding these participants showed that the significance of the results remains the same with or without them. Participants' accuracy on the comprehension questions was 92%, with accuracy rates being similar in questions following the critical and filler items. Two participants who scored lower than 80% on the questions were excluded from the sample and replaced.

7.2. Results

The critical data for analysis are the reading times for the trigger segments in the *Upper-Bound with 'some'* and the *Lower-Bound with 'some'* condition and the target segments in all three conditions. Trigger segments in the *Upper-Bound with 'only some'* condition were not comparable with the other two conditions because they contained an extra word '*only* some of the Fs' which explicitly encodes the inference under study.

7.3. Trigger reading times

The raw data were treated in the same way as in the previous experiments. The exclusion value was calculated as 2500 ms. Over all participants, 3.3% of the data were

removed and smoothed for the *Upper-Bound with 'some'* condition and 2.7% for the *Lower-Bound with 'some'* condition. The mean trigger segment reading times are shown in the top row of Table 5.

Analyses of variance were conducted in which subjects and items were treated as random factors. These analyses revealed a significant main effect of Condition, $F1(1,29)=5.29$, $P<0.05$;. $F2$ $(1,15)=7.81$, $P<0.05$. There was no reliable difference between any other comparable segments of the texts, including the segments following the trigger phrase.

The findings are consistent with Experiments 1 and 2 and support Context-Driven models. Reading time on the implicature trigger ('some of the Fs') was significantly slower in the upper-bound context than the lower-bound context. The results suggest that the implicature was generated in the upper-bound context, but not in the lower-bound context.

7.4. Target reading times

The exclusion value was 1500 ms. Over all participants 2.2% of the data were removed and smoothed for the *Upper-Bound with 'some'* condition, 2.7% for the *Upper-Bound with 'only some'* condition and 3.33% for the *Lower-Bound with 'some'* condition. The mean target segment reading times in these conditions are shown in the bottom row of Table 5.

Analyses of variance were conducted in which subjects and items were treated as random factors. These analyses revealed a significant main effect of Condition, $F1(2,24)=5.03$, $P<0.05$; $F2(2, 30)=4.72$, $P<0.05$. There was no reliable difference between any other comparable segments of the texts, including the segments following the target phrase.

Planned comparisons showed that there was no significant difference between the *Upper-Bound with 'some'* and the *Upper-Bound with 'only some'* contexts, *F1* and *F2* both $<$ 1.0. There was a significant difference between *Lower-Bound with 'some'* and *Upper-Bound with 'some'* contexts, $F1(1,28)=12.45$, $P<0.001$; $F2(1,15)=4.67$, $P<$ 0.05. There was also a significant difference between *Lower-Bound with 'some'* and *Upper-Bound with 'only some'* contexts, $F1(1,28)=9.23$, $P<0.01$; $F2(1,15)=9.0$, $P<$ 0.01. Thus, reading time on the target phrase 'the rest/the other(s)' indicates that the inference was generated in the Upper-Bound context with 'some' and 'only some' but not in the Lower-Bound context with 'some'. The results again support the Context-Driven approach.

7.5. Discussion

This study indicates that a scalar implicature can be generated in sentence final position in Greek, but only when contextual expectations explicitly bias it. It once again seems that the Context-Driven approach is best suited to account for the generation of these implicatures. The Revised Default model, which we discussed at the beginning of this section, predicted that the implicature should not be generated in any of the *with 'some'* conditions since the trigger was in a sentence-final position.

It is worth noting a significant difference between Experiments 2 and 3 regarding the predictions of the standard Default model. In Experiment 2, the sentence-final trigger was contained in a genuinely neutral context while in Experiment 3 it was contained in upper- and lower-bound contexts. According to the standard default model, no implicature is available in lower-bound contexts and so it correctly predicts, along with the Context-Driven model, the difference on the targets between the *Upper-Bound with 'some'* condition and the *Lower-Bound with 'some'* condition as we observed in Experiment 3 (although as we have seen, Default models predict no such difference on the targets in neutral contexts in Experiment 2 – contrary to what was found). However, the standard Default model fails to make the correct predictions in Experiment 3 regarding the triggers. It would predict a negative or no difference between the upper- and lower-bound conditions, whereas what is obtained is a positive difference. This result corroborates that of Experiment 1.

The difference on the target found in Experiment 3 between the *with-'some'* conditions (upper- and lower-bound) also helps us further to factor out any possibility of an effect to do with the explicit control 'only' in Experiment 2. In Experiment 2, we attributed the difference on the target between the *Sentence final-'only some'* and *Sentence final-'some'* condition to the absence of any scalar implicature in the *Sentence final-'some'* condition. Moreover, in the off-line follow-up to Experiment 2, we established that neither of the phrases 'only some of the Fs' nor 'some of the Fs' in sentence final position was significantly preferred over the other in continuations. However, it could be argued that while the phrases both with 'only' and without were the preferred source for continuations in the second sentence, the presence of 'only' still serves to make the complement set (the Fs which did not G) somehow more salient - in spite of the fact that the implicature is generated in the 'some' case. The results from the two upper-bound conditions in Experiment 3 suggest that this line of thinking would be incorrect. In *upper-bound with-'some'* conditions, where it is agreed that an implicature is generated, and where the reading time on the trigger confirms it, there is no difference in response for the target compared to the *with-'only some'* condition.

So the results of Experiment 3 serve to strengthen the grounds for the conclusion invited by the results in Experiment 2: when there is a difference on the target between an *'only some'* and *'some' sentence-final* condition, this difference can be attributed to a lack of implicature generation. Note that since there is a 1000 ms inter stimulus interval between the trigger and the target in Experiment 2, any effects obtained on the target can be straightforwardly attributed to operations taking place on that segment, rather than spill-over effects carrying over from preceding context.

Finally, our results regarding the response time for the target phrases, 'the rest', may be of relevance to research done on the effect of positive and negative quantifiers on discourse focus and anaphora. The general finding (see Moxey, Sanford, & Dawydiak, 2001) is that negative quantifiers (e.g. 'few of the Fs') raise the salience of the complement set (the reference of 'the rest' in our experiments) while positive quantifiers ('some of the Fs') raise the salience of the reference set. Our findings suggest that where scalar implicature is involved, the salience of the complement set is raised.

8. General discussion

In this paper, we set out to delineate the debate between the Context-Driven and the Default approaches to implicature generation by translating their theoretical claims into testable psycholinguistic processing predictions. Our research is motivated by the explicit claims of recent accounts that their predictions should enjoy both linguistic and psycholinguistic validity (Chierchia, 2004; Levinson, 2000; Wilson & Sperber, 2003). Three experiments we reported consistently suggest that the two versions of the default approach we spelled out both over- and under-generate implicatures.

In Experiment 1, we replicated Bezuidenhout and Cutting (2002) study using unequivocal scalar implicature triggers and we found evidence that implicatures are only generated where explicitly warranted by the context. In contexts where the scalar implicature is not warranted because 'all that is relevant is the lower-bound', it is not the case that implicatures are originally generated and subsequently blocked, as the default approach postulated. These findings corroborate the results obtained by Noveck and Posada (2003); Bott and Noveck (2004) through single-sentence truth value judgements tasks. All of these results confirm also the Context-Driven expectation that where implicatures are generated, this takes extra processing resources.

In Experiment 2, we devised materials which both Default and Context-Driven theorists would agree were sufficiently 'neutral'. The evidence suggests that, in the absence of any contextual information about whether or not the upper or lower-bound is relevant, implicatures are not generated. Only when we manipulated sentence position to create implicit contextual expectations which would warrant the implicature did we find evidence for generation.

In Experiment 3 we showed that the fact that implicatures were not generated in sentence final positions in the second experiment is not due to any implicature 'plug' that a revised default model might postulate. We showed that implicatures can be generated even in sentence-final positions, given that readers' expectations are manipulated through explicit contextual expectations. By taking reading time on the implicature triggers as well, Experiment 3 functioned as a replication of Experiment 1, and, in agreement with Experiment 1, we obtained evidence yet again that disconfirms the standard Default model.

All our findings are consistent with the Context-Driven view of language interpretation, where implicatures are processed by a single context-sensitive pragmatic system that cannot be subsumed into the domain of grammar and does not operate on default rules. The pragmatic system would be responsible for the on-line generation of all implicatures, including common scalar implicatures and other so-called GCIs.

The idea that common scalar implicatures are generated by an independent pragmatic system gains some support from studies of child language. What seems quite clear is that children below a certain age do not interpret 'some of the F's' as bearing quantity implicatures even in clear upper-bound contexts (Chierchia, Crain, Guasti, Gualmini, & Meroni, 2001; Crain, Gualmini, & Meroni 2000; Gualmini 2003, Noveck, 2001; Papafragou & Musolino, 2003, i.a.). On the face of it, the developmental delay is difficult to explain on Chierchia's grammatical Default view. However, this data may not be so decisive. There may be a good explanation for the delay relating to independent

performance factors. In particular, on Chierchia's account children still need to compute the alternative interpretation ('all of the Fs have G', where the stimulus is "some of the Fs have G"). The observed delay could be due to children's difficulty with computing these alternatives (see Chierchia et al., 2001).

Although the Context-Driven view does gain support from our study, the results obtained so far are also consistent with an interactive account. Such an account would allocate a circumscribed role for structural factors in addition to contextual factors. Interactive accounts have been explored in depth in the sentence-processing literature. There is a wide range of studies that suggest that contextual assumptions are but one of a set of factors involved in the very first stages of sentence processing. A traditional issue in that literature is whether there is an initial stage of processing where an encapsulated 'privileged' type of information (usually structure-based) operates first, and other types of information are considered later on, potentially causing back-tracking and re-analysis, a phenomenon known as garden-path (Ferreira & Clifton, 1986; Frazier & Rayner, 1982 i.a.). The alternative view is that different types of information, including information from syntax, semantics, linguistic context, the situation referred to and frequency of co-occurrence, interact and coordinate from the earliest possible stage. A large number of studies have suggested that information available in the context is in fact one of the constraints that can affect parsing (Altmann & Steedman, 1988; Spivey-Knowlton & Sedivy, 1995; Taraban & McClelland, 1988; Trueswell, 1996) without it being the only relevant factor. For example, when presenting sentences out of context, there may be a structure-based preference to interpret the ambiguous prepositional phrase 'with the X' as an Instrument rather than as a modifying noun phrase in sentences like (13a,b), which explains our intuition about a garden path in (13b):

13a. John hit the thief with the stick
  b. John hit the thief with the scar

However, when participants are given a context where two referents are salient (for example, two thieves, only one of them with a scar) it has been found that there is no evidence of a garden path on the final phrase (provided that no other constraint is strongly biased).

In the light of this literature, the Default view of GCIs seems to line up with the autonomous, structure-driven view. However, it is not entirely clear that the Context-Driven view of Sperber and Wilson and Carston is closely comparable with the interactionist view. Sperber and Wilson and Carston subscribe to a modular account of language processing according to which grammatically driven semantic representations feed into interpretation processes governed by a separate pragmatics module. Moreover, Sperber and Wilson (2002) are more or less explicit that implicature generation is the product of effortful, attentional processes operating from the earliest stages of language processing. In short, this kind of Context-Driven account would argue that no (non-conventional) implicatures are made available through structure-driven processes.

By contrast, a constraint-based account would allow that not only contextual factors can trigger implicature generation but that certain structural factors could suffice to trigger an implicature. Exactly which factors would serve to trigger the implicature is a question

which constraint-based research would need to investigate. Note however that these structural factors were not necessarily manifest in the studies we have reported here. Such a constraint-based account would also contrast with the Default model we have examined in this paper as it does not predict generation where there are no triggering factors and would not need to rely on cancellation.

We suggest that choosing among interactionist and Context-Driven views is a promising avenue of future research in the study of the pragmatic dimension of language processing. On the face of it, both approaches have their merits and both face challenges. For any interactionist view, there is a burden to come up with a set of context-independent constraints. For the Context-Driven view, there is an issue concerning the fact that some contrastive contextual inferences can affect sentence processing at a very early stage (see Grodner and Sedivy, 2004). While this fact may seem consistent with the Context-Driven view, something needs to be said about how one is to reconcile the earliness of these contrastive inferences with the apparent fact that their nonce derivation would be as complex and as resource-consuming as scalar implicature-described in Section 2 above.

Overall, we have shown that contextual assumptions are taken into consideration when generating scalar implicatures. We have found no evidence in our on-line studies that there is a stage of processing which is impenetrable to contextual assumptions, as a Default position would postulate. We suggest that, in terms of the debate as it currently stands in the pragmatic literature, the encapsulated view of GCI generation is undermined by the experimental evidence reviewed and presented in this paper. However, we have argued that there should be further investigation into whether, as Context-Driven proponents suggest, implicature generation is purely the result of high-level context-dependent computations of the speaker's intended meaning from the earliest stages or whether a constraint-based interactive model would better account for the generation of common conversational implicatures.

Appendix.A. Sample experimental items and comprehension questions

(Original items are in Greek. We present them here in English for ease of exposition)

A.1. Experiment 1: Replicating Bezuidenhout and Cutting

1(a) Upper-Bound: John was taking a university course/and working at the same time./For the exams/he had to study/from short and comprehensive sources./ Depending on the course,/he decided to read/the class notes or the summary./
1(b) Lower-Bound: John heard that/the textbook for Geophysics/was very advanced./ Nobody understood it properly./He heard that/if he wanted to pass the course/he should read/the class notes or the summary./
2(a) Upper-Bound: The day's offer usually is:/You can have a full menu/for one person/and the second person/can have the plat de jour/for free./Today,/customers could have for free/meat or fish./
2(b) Lower-Bound: The dietician that visited the school/explained to children/how useful for our body/protein can be./He also told them/that we can find protein in/meat or fish./

3(a) Upper-Bound: Mary was saying that,/John is so careless,/that he constantly loses things./He has lost money,/keys,/credit cards. Once/along with his bag,/he even managed to loose/his ID or his passport./

Comprehension question: Are credit cards among the things that John has lost?

3(b) Lower-Bound: The police stopped John/for a routine control/but John had only/his driving license on him./The policemen/stressed to John that valid/identification documents are/his ID or his passport./

Comprehension question: Was it a driving license that John had on him?

4(a) Upper-Bound: While Mary and John were out shopping,/it started raining./John would get wet./Even though she did not have a lot of money,/she offered to buy him/an umbrella or a coat./

4(b) Lower-Bound: It was highly probable that it would rain./Mary advised John/to dress accordingly./To avoid getting wet,/she suggested to him/to take with him/an umbrella or a coat./

A.2. Experiment 2: scalar implicature generation in neutral contexts

1(a) Sentence initial-'some': Some of the consultants/had a meeting/with the manager./The rest/did not manage/to attend./

1(b) Sentence initial-'only some': Only some of the consultants/had a meeting/with the manager./The rest/did not manage/to attend./

1(c) Sentence final-'some': The manager/had a meeting/with some of the consultants./The rest/did not manage/to attend./

1(d) Sentence final-'only some': The manager/had a meeting/with only some of the consultants./The rest/did not manage/to attend./

2(a) Sentence initial-'some': Some of the tourists/got soaked/by the rain storm/. The others/had taken care to bring an umbrella/just in case.

2(b) Sentence initial-'only some': Only some of the tourists/got soaked/by the rain storm/. The others/had taken care to bring an umbrella/just in case.

2(c) Sentence final-'some': The rain soaked/some of the tourists. The others/had taken care to bring an umbrella/just in case.

2(d) Sentence final-'only some': The rain soaked/only some of the tourists. The others/had taken care to bring an umbrella/just in case.

3(a) Sentence initial-'some': Some of the boys/went to the cinema/. The others/went for bowling/and roller-skating.

3(b) Sentence initial-'only some': Only some of the boys/went to the cinema/. The others/went for bowling/and roller-skating.

3(c) Sentence final-'some': John went to the cinema/with some of the boys/. The others/went for bowling/and roller-skating.

3(d) Sentence final-'only some': John went to the cinema/with only some of the boys/. The others/went for bowling/and roller-skating.

  Comprehension question: Did any of the boys go to the theatre?

4(a) Sentence initial-'some': Some of the jokes/were funny/. The others/were rather too corny/and stereotypical.
4(b) Sentence initial-'only some': Only some of the jokes/were funny/. The others/were rather too corny/and stereotypical.
4(c) Sentence final-'some': Mary laughed/at some of the jokes/. The others/were rather too corny/and stereotypical.
4(d) Sentence final-'only some': Mary laughed/at only some of the jokes/. The others/were rather too corny/and stereotypical.

A.3. Experiment 3: Controlling for plugs

1(a) Implicit Upper-Bound: Mary asked John/whether he intended to host/all his relatives/in his tiny apartment./John replied/that he intended to host/some of his relatives./The rest/would stay/in a nearby hotel./
1(b) Explicit Upper-Bound Context: Mary asked John/whether he intended to host/all his relatives/in his tiny apartment./John replied/that he intended to host/only some of his relatives./The rest/would stay/in a nearby hotel./
1(c) Implicit Lower-Bound Context: Mary was surprised/to see John/cleaning his apartment/and she asked/the reason why./John replied/that he intended to host/some of his relatives./The rest/would stay/in a nearby hotel./

  Comprehension question: Was it his schoolmates that John intended to host?

2(a) Implicit Upper-Bound: Mary wanted to create/a very special atmosphere/and wondered whether to light all her special candles/in the dining hall or not./Eventually she took her decision:/she lit/some of her candles./The others/were left unlit/to create a nice atmosphere.
2(b) Explicit Upper-Bound Context: Mary wanted to create/a very special atmosphere/and wondered whether to light all her special candles/in the dining hall or not./Eventually she took her decision:/she lit/only some of her candles./The others/were left unlit/to create a nice atmosphere.
2(c) Implicit Lower-Bound Context: Mary pays attention/to the most minute detail/in order to create a romantic atmosphere/in her dinner parties./Last night she had a dinner party and/she lit/some of her candles./The others/were left unlit/to create a nice atmosphere.
3(a) Implicit Upper-Bound: Yesterday, George asked Mary/whether John had gone to the cinema/with all his friends/. Mary said that/he went to the cinema/with some of his friends./The others/went for bowling/and roller-skating.

3(b) Explicit Upper-Bound: Yesterday, George asked Mary/whether John had gone to the cinema/with all his friends/. Mary said that/he went to the cinema/with only some of his friends./The others/went for bowling/and roller-skating.

3(c) Implicit Lower-Bound Context: John's mother/is anxious about his where-abouts./She asked Mary/if she knew what John did on Friday night./Mary said that/he went to the cinema/with some of his friends./The others/went for bowling/and roller-skating.

4(a) Implicit Upper-Bound: The warden of the zoo/asked whether all the lions had been fed./The worker replied that/he had fed/some of the lions./The rest/were being examined by the vets/and would be fed later.

4(b) Explicit Upper-Bound: The warden of the zoo/asked whether all the lions had been fed./The worker replied that/he had fed/only some of the lions./The rest/were being examined by the vets/and would be fed later.

4(c) Implicit Lower-Bound Context: The warden of the zoo/asked the worker/what he had been doing the whole morning./He replied that/he had fed/some of the lions./The rest/were being examined by the vets/and would be fed later.

References

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *38*, 419–439.

Bezuidenhout, A., & Cutting, J. C. (2002). Literal meaning, minimal propositions, and pragmatic processing. *Journal of Pragmatics*, *34*, 433–456.

Bezuidenhout, A., & Morris, R. (2004). Implicature, relevance and default inferences. In I. Noveck, & D. Sperber (Eds.), *Experimental pragmatics*. Basingstoke: Palgrave Macmillan.

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

Carston, R. (1998). Informativeness, relevance and scalar implicature. In R. Carston, & S. Uchida (Eds.), *Relevance theory: Applications and implications*, 179–236.

Carston, R. (2002). *Thoughts and utterances*. Oxford: Blackwells.

Carston, R. (2004). Review of Stephen C. Levinson presumptive meanings: The theory of generalized conversational implicature. *Journal of Linguistics*, *40*, 181–186.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (Ed.), *Belletti structures and beyond*. Oxford: Oxford University Press.

Chierchia, G., Crain, S., Guasti, M., Gualmini, A., & Meroni, L. (2001). *The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures Proceedings from the 25th annual BUCLD*. Somerville, MA: Cascadilla Press pp. 157–168.

Crain, S., Gualmini, A., & Meroni, L. (2000). The acquisition of logical words. *LOGOS and Language*, *1*, 49–59.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348–358.

Frazier, L., & Clifton, C. (1996). *Construal*. Cambridge, MA: MIT Press.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*, 291–325.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*, 178–210.

Gazdar, G. (1979). *Pragmatics: Implicature, presupposition and logical form*. New York: Academic Press.

Geurts, B. (1998a). Scalars. In P. Ludewig, & B. Geurts (Eds.), *Lexikalische Semantik aus kognitiver Sicht* (pp. 95–118). Tübingen: Gunter Narr.

Geurts, B. (1998). Presuppositions and anaphors in attitude contexts. In *Linguistics and Philosophy* Vol. 21, pp. 545–601.

Gibbs, R., & Moise, J. (1997). Pragmatics in understanding what is said. *Cognition*, *62*, 51–74.

Grice, H.P. (1989). Logic and conversation. In Studies in the way of words (pp. 22–40). Cambridge, MA: Harvard University Press.

Grodner, D., & Sedivy, J. (2004). The effect of speaker-specific information on pragmatic inferences. In N. Pearlmutter, & E. Gibson (Eds.), *The processing and acquisition of reference*. Cambridge, MA: MIT Press.

Hobbs, J., Stickel, M., Appelt, D., & Martin, P. (1993). Martin interpretation as abduction. *Artificial Intelligence*, *63*, 69–142.

Horn, L. (1984). Toward a new taxonomy for pragmatic inference. In D. Schiffrin (Ed.), *Form and use in context: Linguistic applications (GURT '84)* (pp. 11–42). Washington, DC: Georgetown University Press.

Horn, L. (1992). The said and the unsaid. Ohio state university working papers in linguistics (SALT II proceedings), Vol. 40 (pp. 163–192). Ohio State University.

Katsos, N., Breheny, R., & Williams, J. (2005). *Interaction of structural and contextual constraints. During the on-line generation of scalar inferences Proceedings of GLOW 28, University of Geneva*.

Levinson, S. (2000). *Presumptive meanings*. MIT Press: Cambridge, MA.

Moxey, L. M., Sanford, A. J., & Dawydiak, E. J. (2001). Denials as controllers of negative quantifier focus. *Journal of Memory and Language*, *44*(3), 427–442.

Nicolle, S., & Clark, B. (1999). Experimental pragmatics and what is said. *Cognition*, *69*, 337–354.

Noveck, I. (2001). When children are more logical than adults. *Cognition*, *86*, 253–282.

Noveck, I., & Posada, A. (2003). Characterising the time course of an implicature. *Brain and Language*, *85*, 203–210.

O'Brien, E. J., Duffy, S. A., & Myers, J. L. (1986). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language*, *27*, 399–415.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics/pragmatics interface. *Cognition*, *86*, 253–282.

Philippaki-Warburton, I. (1987). The theory of empty categories and the pro-drop parameter in Modern Greek. *Journal of Linguistics*, *23*, 289–318.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.

Sperber, D., & Wilson, D. (2002). Pragmatics modularity and mind reading. *Mind and Language*, *17*, 3–23.

Spivey-Knowlton, M., & Sedivy, J. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, *55*, 227–267.

Stanley, J. (2000). Context and logical form. *Linguistics and Philosophy*, *23*, 391–434.

Taraban, R., & McClelland, J. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content based expectations. *Journal of Memory and Language*, *27*, 1–36.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, *35*, 566–585.

van Kuppevelt, J. (1996). Inferring from topics. *Linguistics and Philosophy*, *19*, 393–443.

Westerstahl, D. (1985). Determiners and context sets. In J. van Benthem, & A. ter Meulan (Eds.), *Generalised quantifiers and natural language* (pp. 45–71). Dordrecht: Foris.

Wilson, D., & Sperber, D. (2003). Relevance theory. In L. Horn, & G. Ward (Eds.), *Handbook of pragmatics*. Oxford: Blackwell pp. 607–632.