

Interaction of Structural and Contextual Constraints During the On-line Generation of Scalar Inferences

Napoleon Katsos (nk248@cam.ac.uk)

RCEAL, University of Cambridge, 9 West Road, CB3 9DP, UK

Richard Breheny (richardb@linguistics.ucl.ac.uk)

Department of Phonetics and Linguistics, University College London
Gower Street, WC1E 6BT UK

John Williams (jnw12@cam.ac.uk)

RCEAL, University of Cambridge, 9 West Road, CB3 9DP, UK

Abstract

We investigate the interaction of structural and contextual constraints on the on-line generation of three types of Scalar Implicature, (a) the disjunction: “A or B” >> “either A or B but not both”, (b) the partitive NP: “some of the Fs” >> “at least one but not all of the Fs” and (c) ad hoc context dependent scales. In theoretical linguistics, according to the structural approach (Chierchia, 2004; Levinson 2000) Scalar Implicatures (SIs) are generated whenever certain constraints on the semantic properties of the linguistic structure are satisfied; SIs have to be cancelled subsequently if information from the discourse context indicates that the SI is not felicitous. According to the pragmatic approach, structural properties are a necessary but not sufficient condition, since for SIs to be generated there ought to be further contextual constraints that make the inference relevant (Recanati, 2003; Sperber & Wilson, 1995). The linguistic debate is reflected in the psycholinguistic discussion on the modular versus interactive nature of the human parser. The structural approach presupposes a modular and serial parser and the pragmatic an interactive one which allows multiple constraints to operate from the earliest stage. Recent psycholinguistic studies on SIs are not reaching a consensus. Single sentence truth value judgment tasks (Bott & Noveck, 2004; Noveck & Posada, 2003) support the pragmatic/interactionist approach, but text comprehension studies (Bezuidenhout & Cutting, 2002; Bezuidenhout & Morris, 2004) have been ambivalent and studies on the visual-world eye-tracking paradigm (Storto, Tanenhaus & Carlson, 2005; Huan & Snedeker, 2005) are addressing only some of the critical issues. We present one off-line and three on-line text comprehension experiments that conclude in favor of the pragmatic/interactionist approach. We discuss the relevance of the findings for psycholinguistics and linguistics.

Introduction

Certain linguistic expressions form entailment scales where terms on the right of the scale are informationally stronger than terms on the left (measured by number of entailments, e.g. <some, most, all>, <or, and>). Scalar expressions may trigger Scalar Implicatures (SIs) in upward entailing structures when the assertion of an informationally weaker term conversationally implies (“>>”) the negation of the stronger terms in (1).

- (1) Mary: How many students failed the test?
Jane: Some of them did >> Not all of them did

A distinctive property of SIs is defeasibility due to structural and contextual constraints. Regarding the former, SIs are dependent on the direction of the entailment of the scale. In downward entailment structures, e.g. in the antecedent of conditionals, the direction of entailment in the scale is reversed (“some” entails “all”) and SIs are not generated:

- (2) If some of the students fail the test, their teacher will be disappointed.

Here there is no implication that if all the students fail the test their teacher will not be disappointed.

Moreover, SIs are context dependent. There are three types of relevant contexts, (a) upper bound, where it is necessary to draw a SI in order to satisfy a discourse goal like in (1), (b) neutral contexts, where there is no assumption as to whether the SI is relevant or not, and (c) lower bound contexts, where the discourse goal can be satisfied by the plain meaning of the SI trigger without any need for the inference. It is assumed that SIs are not generated in lower bound contexts since the assertion of stronger terms of the scale is irrelevant:

- (3) Mary: Why is the teacher disappointed?
Jane: Some of his students failed the test.

In (3) the discourse goal that is set by Mary’s question is satisfactorily addressed by the plain meaning of “some”, “at least one of the students and even all”, so it becomes irrelevant whether the strengthened alternative “at least one but not all” is true or not (e.g. Levinson, 2000).

The linguistic debate

The current linguistic debate is set between the structural approach (Chierchia, 2004; Levinson, 2000) which posits that logical properties of the linguistic structure are the necessary and sufficient conditions for SI generation, and the pragmatic approach (Carston, 1998; Recanati, 2003;

Sperber & Wilson, 1995) that posits an interaction between the three types of context (upper-, lower-bound and neutral) and structural constraints.

The structural approach

According to Chierchia’s version, SIs are introduced locally, below the level of the full sentence, and automatically, as soon as a scalar term enters the computation. The grammar assigns two values for the scalar term, the plain meaning of the expression and its scalar alternative which incorporates the SI. Subsequently, the informationally weaker meaning will be filtered out. In the case of upward entailing structures this is the plain meaning of the term, but in downward entailing structures this is the meaning with the SI. Thus, there are cases where SIs are computed by default and cases where they are absent by default; in both cases the key factor is structural properties. At a second stage, the SI is checked against the background assumptions. If the addition of a default SI to a given discourse context is infelicitous, (e.g. to a lower bound context as in 3), the SI will be cancelled; but this happens only at a later stage after compositional grammatical processes have been concluded. A corollary of the structural approach is that SIs are linguistic inferences generated by a default process that should come for little processing cost.

The pragmatic approach

According to the pragmatic approach the structural constraint of upward entailment is indeed a necessary condition for SI generation and SIs will not be generated in downward entailing structures. However, for SIs to be generated, besides being in an upward entailing structure, they must also be under an upper-bound discourse context, i.e. a discourse goal which makes the SI relevant (e.g. for the SI term “some of the Fs” this would be a discourse where the quantity of the Fs involved is important).

The differences between the structural and the pragmatic approach become evident in upward entailing structures (the structures that allow for SI generation) in two respects: First, given lower-bound discourse contexts where the SI is incompatible with the discourse, the first predicts that SIs will be generated and subsequently cancelled and the second that SI will not be generated at all. Moreover, in upward entailing structures which are discourse neutral (i.e. there are no contextual assumptions that make the SI relevant), the structural approach predicts the generation of SIs regardless of the absence of contextual assumptions whereas the pragmatic approach predicts that the SI will not be generated. Table 1 summarizes the predictions:

Table (1). Structural & Pragmatic Approach Predictions

	Structural Approach	Pragmatic Approach
(i) SIs generated locally	YES	YES
(ii) SIs are not generated in downward entailment structures	YES	YES

(iii) In upward entailing structures with an upper-bound discourse context	SI generated	SI generated
(iv) In upward entailing structures which are discourse neutral	SI generated	SI NOT generated
(v) In upward entailing structures with a lower-bound discourse context	SI generated & cancelled	SI NOT generated
(vi) In terms of processing resources	generation costless & cancellation costly	generation costly & cancellation never occurs

Chierchia (2005), Levinson (2000), Sperber & Wilson (1995) and other linguists have explicated the need for relevant psycholinguistic evidence.

Psycholinguistic investigation

The linguistic debate is related to research in psycholinguistics on whether there is an initial stage of processing where an encapsulated 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 & Clifton, 1996 i.a.). The alternative view is that different types of information from syntax, semantics, linguistic context, the situation referred to and statistical frequency of co-occurrence among other, interact and coordinate from the earliest possible stage.

Though the debate is not resolved, a large number of studies have shown that information available from the context is in fact one of the constraints that can affect parsing (Altmann & Steedman, 1988; Spivey-Knowlton. & Sedivy, 1995 i.a.) though the timing in which these constraints operate need not be the same (Gibson & Pearlmutter, 1998). A serial and modular parser assuming that information from the grammar governs the initial stages of parsing would be expressed though a version of the structural approach in linguistics. Interactive models would be expressed through a version of the pragmatic approach. Such models accept that a number of constraints may guide the first stages of parsing and could allow for pragmatic factors such as discourse goals to be one type of such constraints. In fact the work of Sedivy and colleagues (e.g. Grodner & Sedivy, in press; Sedivy et al, 1999) on the visual-world eye-tracking paradigm has shown that pragmatic and discourse factors similar to the ones that we discuss can affect parsing from the earliest possible stage. There is growing body of experimental research directly focused on SI processing which we will discuss.

Text comprehension

The study reported in Bezuidenhout & Cutting (2002) investigates prediction (v), the case where an SI in an upward entailing structure is incompatible with a lower-bound discourse context. Bezuidenhout & Cutting (hence B&C) test various scales including the numerals and

quantifiers. According to the structural approach, in a lower-bound context the SI has to be cancelled, leading to longer (or at least equal) reading time than in an upper-bound context where the SI is allowed to be generated. In a self-paced reading experiment B&C found that there was a significant slowdown when reading the SI triggers in lower-bound contexts compared to their upper-bound versions. They conclude that in lower-bound contexts SIs are indeed first generated and then cancelled, supporting the structural approach.

However, we must note that B&C's experimental items were not "orthodox" implicatures according to the current consensus in linguistic theory. The cardinals by which they exemplify their case are not considered implicatures by most theorists in either camp of the debate (see Breheny, in press; Geurts, 1998; Papafragou & Musolino 2003 for developmental evidence). Other expressions B&C used have never been considered implicatures. E.g. inferring from 'everyone is a vegetarian' that 'everyone in this group (= not in the whole universe) is a vegetarian' is a process of narrowing down the range of application of the quantifier and not of implicature generation. All in all, almost half of the categories B&C tested do not give rise to SIs.

Bezuidenhout & Morris (2004; hence B&M) used an eye-tracker to record reading time and gaze fixations on sentences like (4a) where the 'not all of the Fs' SI is explicitly cancelled in the second sentence by the phrase 'in fact all of them did'. They used definite expressions as a control (4b):

- (4) a. Some books had colour pictures. In fact all of them did, which is why the teachers liked them.
b. The books had colour pictures. In fact all of them did, which is why the teachers liked them.

According to the structural approach, participants should generate the SI when 'some' is read and will engage in cancellation and reanalysis at the 'them did' region of the sentence. In contrast, according to B&M, the pragmatic approach claims that participants do not fully commit to the 'some but not all' SI right away. The pragmatic approach predicts increased processing time on the word 'all' to reflect the fact that the reader has registered information potentially relevant to the specification of an underspecified item.

B&M find that there is no reading time slowdown at the 'them did' region between (4a) and (4b) but there is a slowdown on 'all' in (4a), which they interpret as evidence in favor of the pragmatic approach. No differences were found in gaze fixations. There are however a number of objections. Firstly, the discourses used were not appropriate since a) at the beginning of the second phrase there always was a discourse marker which indicated contrast ('in fact', 'actually' etc) which is not neutral to the inferences under study and b) their control example (4b) is an infelicitous discourse and further questions arise from a follow up study that was designed to control for this. Second, B&M's understanding of underspecification is problematic, implying that interpretations can remain underspecified indefinitely in case there is no further relevant information

in the discourse. It is our understanding that both structural and pragmatic approaches clearly ascribe to the prediction that if an SI is generated then this should occur immediately once the trigger has been processed. The studies by B&C and B&M have marked the beginning of SI experimental investigation but for different reasons both studies remain inconclusive.

Single sentence truth value judgment tasks

When presented with an SI term and asked to make a judgment, respondents may answer "false" to a sentence like (5) if they generate the SI ("at least one but not all") or answer "true" if they use the plain meaning ("at least one and perhaps all"):

- (5) Some elephants have trunks

Noveck & Posada (2003) instructed groups of participants to use scalar terms either with the SI or just with the plain meaning. They timed the responses in a sentence verification task and found that the time taken by the participants answering 'false' was significantly longer than those answering 'true'. That is, when respondents based their answers on the plain meaning, it was not the case that the SI was first generated and then cancelled but rather that the SI was not generated in the first place. This result supports a model where SIs are costly inferences. In a similar study Bott & Noveck (2004) replicated these findings and also excluded the possibility that the difference was due to the difference in response type by including appropriate control conditions. Moreover, in a second experiment they introduced an additional layer of narration, where the sentences are attributed to a character. The stimuli are preceded either by the declaration that 'Mary says the following sentence is true' or that 'Mary says the following sentence is false'. Thus they replicated the previous findings with a design where it is possible for answers based either on the SI or the plain meaning to have the same response type. In the third experiment they removed the instructions to the participants group and still obtained the same pattern and in the fourth experiment they found that the number of responses based on the SI increased as permitted response time increased. The studies by Noveck and colleagues have yielded consistent evidence against the structural approach regarding prediction (v) and prediction (vi).

Visual-world eye-tracking

In a recent study Storto, Tanenhaus & Carlson (2005) investigated the scale <or, and>, whereby the plain meaning of the disjunction "or", called the inclusive interpretation, is "A or B or even both" and the SI gives rise to the exclusive interpretation "either A or B but not both". Storto et al. found that the SI was calculated and integrated very locally to the utterance of the disjunction and could guide the further processing of the sentence. Thus they supported prediction (i), that the process that generates SIs operates fast, below the level of the full sentence, which is shared by both the structural and the pragmatic approach. However, in their study there is no lower-bound condition, where the disjunction would have an inclusive interpretation which

would allow testing prediction (v) where the two approaches make differential claims. Huang & Snedeker (2005) compared the processing of number terms and the partitive noun phrase “some of the F” using a similar paradigm. They found evidence that supports the view that numerals are categorically different from scalar terms proper. When presented with a set of plausible alternatives (one, two or three Fs) subjects in the number condition fixated on the correct image from the beginning and engaged in very early disambiguation of the sentence. In contrast, subjects in the scalar condition developed this preference significantly later and initially looked at all plausible alternatives (“some of the Fs” and “all the Fs”), indicating relatively late disambiguation. Besides making a claim on the nature of the numerals, these findings suggest that even though SIs are generated on-line locally, the process that generates them is more time-consuming than the process that generates grammar driven interpretations like the numerals.

The present research

We aimed to devise experiments that would put to test all six predictions of the structural and the pragmatic approach. Since the findings from the single sentences judgment tasks are conclusive, we focused on SIs that were embedded in a discourse and tried to overcome the shortcomings of the previous studies.

An off-line study on defeasibility. In our first experiment, we tested the assumption of both the structural and the pragmatic approach that SIs do not arise (a) in upward entailing structures with lower-bound contexts and (b) in a downward entailing structures. Even though this is a critical assumption, according to our knowledge it has not been experimentally addressed up to now. We tested the SI associated with the exclusive interpretation of the disjunction “A or B” >> “either A or B but not both”. We created pairs of sentences where the first sentence was either a question which requires a single entity for an answer, therefore creating an upper-bound context, or a question which can be answered by a list of entities, creating a lower-bound context. In the upper- and lower-bound condition the second sentence was identical and contained a disjunction in an upward entailing structure that answered the preceding question. In the downward entailing condition the question in the first sentence was upper-bound, but the second sentence embedded the SI trigger in the antecedent of a conditional which is a downward entailing structure. We ran an off-line sentence completion task where the final verb of the second sentence was missing, and participants were asked to fill it in using a closed class of verbs (is/are, does/do, has/have) marked for number. Our assumption is that if participants generate the SI that “not both A or B did so” they will use a verb form to agree with the noun phrase “A or B” in singular number. If they do not generate the SI but use the plain meaning of the disjunction “A or B or even both”, they will be much more prone to use a verb that will agree with the disjunction in plural number.

Design and results. We created 18 critical items each in 3 conditions (upper-bound, UB, lower-bound, LB, and

downward entailing, DE, (6)) which were rotated in 3 presentation lists (total n of participants = 45, speakers of British English). 18 fillers were added which had a similar structure with the critical items, but at the place of the SI trigger they contained one proper name (e.g. ‘Jones’).

- (6) **UB:** The manager asked: Who has the report on last year’s profits? Her secretary replied: Jones or Barnes from the department of Finance...
LB: The manager asked: Who has a report on last year’s profits to show me? Her secretary replied: Jones or Barnes from the department of Finance...
DE: The manager asked: Who has the report on last year’s profits? Her secretary replied: If Jones or Barnes ... it, I will bring the report on your desk in a few minutes.

Participants used a verb in singular agreement 82% in upper-bound, 52% in lower-bound and 51% in downward entailing. There was a main effect of Condition ($F(2, 39) = 52.75, p < 0.001$) indicating that participants built an exclusive interpretation in the first case and an inclusive in the latter two (a smaller effect was obtained with speakers of American English). It is suggested that indeed SIs are generated in upward entailing contexts but not generated in downward entailing structures, as both the structural and the pragmatic approach assume in predictions (ii) and (iii). This is also the case in the lower-bound context, but from this study we cannot know whether the end result (no SI) was achieved through the generation and subsequent cancellation of an SI or simply without the generation of the SI in the first place. This issue is addressed in prediction (v).

On-line processing in upper- and lower-bound contexts.

To test prediction (v) we ran an on-line reading time experiment. The structural approach predicts that the SI trigger will be read faster (or at least equal) in the upper-bound than the lower bound context (where the SI has to be generated and then cancelled) and the pragmatic predicts that it will be read slower in the upper-bound (because the SI is not generated at all).

Design and results. We used the same 18 critical items that we employed in experiment 1 in two conditions, the upper- and lower-bound context. We completed the missing verb in the second sentence of each item with the appropriate verb in singular. The items were rotated in 2 presentation lists (total n of participants = 30). 60 items were added which were the critical items of unrelated experiments. Reading times were recorded in a segment by segment self-paced reading paradigm whereby the whole phrase that contained the disjunction appeared as one segment on the screen. Comprehension was motivated by questions.

- (7) **UB:** The manager asked: Who has the report on last year’s profits? Her secretary replied: /Jones or Barnes/ from the department of Finance has. Would you like to see the report?
LB: The manager asked: Who has a report on last year’s profits to show me? Her secretary replied:

/Jones or Barnes/ from the department of Finance has. Would you like to see the report?

The critical segment is presented within brackets. This segment was read in 819ms in the upper-bound and 775ms in the lower bound condition. There was a main effect of Condition, $F(1,38) = 5.05, p < 0.03$. In line with the results of Noveck and colleagues, it is suggested that it wasn't the case that in the upper-bound condition the SI was first generated and then cancelled, but rather that it wasn't computed in the first place. We also see that SI are generated locally, before the end of the sentence (as both approaches expect, in prediction *i*). It is also suggested that computing an SI is actually costly as reflected in reading time required (prediction *vi*).

On-line processing in ad hoc scales. In experiment 3 we replicated the finding that SI generation is a local and a time-consuming process by using non-logical scales introduced by the discourse context (e.g. <roof, house>, <father, parents>). We created 12 items in upper-and lower-bound conditions (8) which were rotated in two presentation lists (total n of participants = 30). Another 50 unrelated items were added as fillers. Comprehension was motivated by questions (the critical segment is presented within brackets).

- (8) **UB:** George went to pick up Mary from the station. He was covered in paint. Mary asked him: Were you painting the house? George replied: I was painting/ the roof/ with an insulating paint.
LB: George went to pick up Mary from the station. He was covered in paint. Mary asked him: What were you painting? George replied: I was painting/ the roof/ with an insulating paint.

In a self-paced reading time paradigm the UB condition was read in 677ms and the LB in 623ms. There was a main effect of Condition ($F(1, 16) = 8.24, p < 0.02$) for the critical segment (and a marginal spillover effect ($p < 0.09$) in the next segment), in line with the findings of experiment 2.

On-line processing in neutral contexts. In a previous study, Breheny, Katsos & Williams (under revision) created discourse neutral contexts by presenting pairs of sentences without preceding context. We manipulated a discourse constraint, the sentence's information structure, and put the SI trigger "some of the Fs" in either Topic or Non-Topic position in the first sentence. We thus introduced implicit assumptions about how relevant the group of Fs is. The second sentences contained the anaphoric expression "the rest of the Fs" whose interpretation is facilitated if participants have inferred that "some of the Fs G-ed" implies that "not all the Fs G-ed". We compared reading time in each condition with a control condition where the SI was made explicit by the operator "only" (which automatically introduces alternative sets on-line (Ni, Crain, & Shankweiler 1996).

Design and Results. 24 critical items were created in 4 conditions (Topic, Topic & Only, Non-Topic, Non-Topic & Only, see 9) and rotated in 4 presentations lists. The experiment was run in Greek, which has a flexible word order and the initial position on the sentence is consistently associated with topicality. 48 participants took part. Comprehension was motivated by questions.

- (9) **T:** Some of the consultants met the manager. /The rest/ did not manage to attend.
T-Only: Only some of the consultants met the manager. /The rest/ did not manage to attend.
Non-T: The manager met some of the consultants. /The rest/ did not manage to attend.
Non-T-Only: The manager met only some of the consultants. /The rest/ did not manage to attend.

The results are shown in table (2):

Table (2). Mean reading time and SD of experiment 4

	Reading time (ms)	SD
Topic	613	125
Topic & Only	611	110
Non-Topic	628	138
Non-Topic & Only	586	112

The critical interaction between position (Topic/Non-Topic) and the presence of Only was significant, by subjects $F_s(1,44) = 4.24, p < 0.05$; and by items $F_i(1,19) = 6.93, p < 0.05$. Reading times were slower in the Non-Topic than the Non-Topic & Only condition, but there was no such difference between the Topic and the Topic & Only conditions. It is suggested that contrary to the structural approach (prediction *iv*), the SI was only generated when the trigger phrase was in one of the topic conditions, even though both the topic and the non-topic positions were in structurally identical sentences (upward entailing).

In this paper, we investigated whether the slow reading times in the Non-Topic condition obtained by Breheny, Katsos & Williams (under revision) could be due to a potential topic shift between the trigger and the target sentence. When participants read the second sentence they might assume that the topic of the first sentence, the first NP mentioned, will be the preferred continuation. The target phrase "the rest" refers to the second NP, thus violating topic continuity. If this potential topic shift is avoided in the case of the Non-Topic & Only due to the presence of "only", then the slow down in the Non-Topic condition may not be reflecting SI generation but topic shift penalty. We ran an off-line sentence continuation task with the items used in experiment 4 with 20 participants. In the Non-Topic condition the continuations referred to NP1: 30%, to NP2: 52,5%, and OTHER: 17,5%. In the Non-Topic & Only to NP1: 24,2%, to NP2: 59,2%, and OTHER: 16,6%. Crucially, both conditions prefer an NP2 continuation over an NP1 ($F_{NP1}(1, 16) = 5.01, p < 0.04$; $F_{NP2}(1, 16) = 12.09, p < 0.003$); their difference in this preference is not significant ($F(1, 16) = 0.022$ n.s.). They are therefore paying an equal topic shift penalty. However, we also ran a second on-line

follow-up study similar to the original one except that the target phrase ‘the rest’ was in non-topic position, avoiding any topic shift penalties:

- (10) **Non-topic (& Only):** The manager met (only) some of the consultants. He will talk/ to the rest/ next week.

The results replicated the original study that SIs are not generated in the Non-Topic condition even though this condition is structurally identical to the Topic one.

Discussion

In this paper we discuss four on- and off-line text comprehension experiments that tested 6 predictions on the processing of Scalar Implicatures. From experiment 1 we provided processing evidence for the fundamental assumptions shared by all approaches on SIs and in experiments 2 to 4 we gathered evidence in favor of the pragmatic/interactionist approach which predicts an interaction between structural properties of the sentence and the type of discourse context. Our findings are aligned with the work of Noveck and colleagues on single sentence truth value judgment task and work on the visual-world eye-tracking paradigm (Huang & Snedeker, 2005). It is suggested that pragmatic considerations is one type of constraint that can affect the earlier stages of processing, supporting an interactive rather than modular type of parser (e.g. Altmann & Steedman, 1988). Our work is also in line with research by Sedivy and colleagues on reference resolution has identified the on-line influence of pragmatic constraints that are similar to the ones we discuss in this paper. A major finding is that however local their generation is, Scalar Implicatures are time-consuming inferences; further research could identify the role of working memory.

Acknowledgments

We would like to thank Anne Bezuidenhout, Robyn Carston, Gennaro Chierchia, Bart Geurts, Ted Gibson, Stavroula Kousta, Ira Noveck, Anna Papafragou, Doug Saddy, and Deirdre Wilson for comments and input. The first author was supported by a Gates Cambridge Scholarship and an E.S.R.C. grant (PTA-030-2002-01744). Many thanks to all the participants of the experiments.

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 Noveck, I. & Sperber, D. (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
- Breheny, R. (forthcoming). Some Scalar Implications Really Aren't Quantity Implicatures - But Some Are. *Proceedings of Sinn & Bedeutung 9*, University of Nijmegen.
- Breheny, R., Katsos, N. & Williams, J. (under revision) Are Generalised Scalar Implicatures Generated by Default? An on-line investigation into the role of context in generating pragmatic inferences.
- Carston, R. (1998). Informativeness, relevance and scalar implicature. In R. Carston & S. Uchida (Eds.), *Relevance Theory: Applications and Implications*. Amsterdam: John Benjamins.
- Chierchia, G. (2004). Scalar implicatures, polarity phenomena and the syntax/pragmatic interface. In A. Beletti (Ed.), *Structures and Beyond*. Oxford University Press.
- 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.
- Geurts, B. (1998). Scalars. In P. Ludewig and B. Geurts (Eds.), *Lexikalische Semantik*. Tübingen: Gunter Narr.
- Gibson, E. & Pearlmutter, N. (1998). Constraints in Sentence Comprehension. *Trends in Cognitive Science*, 2, 262-268
- Grodner, D. & Sedivy, J. (in press) The effect of speaker-specific information on pragmatic inferences. In N. Pearlmutter & E. Gibson (Eds.), *The Processing and Acquisition of Reference*. MIT Press: Cambridge, MA
- Huang, Y.T. & J. Snedeker (2005) What exactly do numerals mean? Paper presented at *Experimental Pragmatics: The cognitive basis of conversation*. Cambridge, UK.
- Levinson, S. (2000). *Presumptive Meanings*. Cambridge, MA.: MIT Press.
- Noveck, I. A. & A. Posada (2003). Characterising the time course of an implicature. *Brain and Language*, 85, 203-210.
- Papafragou, A. & Musolino J. (2003). Scalar implicatures: experiments at the semantics/pragmatics interface. *Cognition*, 86, 253-282.
- Recanati, F. (2003). *Literal Meaning*. Cambridge: Cambridge University Press
- Sedivy, J., Tanenhaus, M., Chambers, C., & Carlson, G. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71, 109-147.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Spivey-Knowlton, M. & Sedivy, J. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, 55, 227-267.
- Storto, L., M. K. Tannenhaus & G. Carlson (2005) The calculation of scalar implicatures: Evidence from eye-tracking. Paper presented at the *Annual Meeting of the Linguistic Society of America*. Oakland, California.