Scalar implicature: Inference, convention, and dual processes *

Keith Frankish[a] and Maria Kasmirli[b]

1. Introduction

An utterance is said to generate an *implicature* if it conveys something that goes beyond what is literally said – if hearers form a pragmatic interpretation of the utterance that differs from the literal one. Common examples are *scalar* implicatures, where a weaker claim is interpreted as the denial of a stronger one on the same scale. For example 'Some politicians take bribes' pragmatically implies 'Not all politicians take bribes', even though 'some' means *at least one* and is logically compatible with 'all'.

The first major analysis of implicature was by the philosopher Paul Grice, who coined the term 'implicature' (see his 'Logic and conversation' reprinted in Grice, 1989). Philosophers of language have continued to analyse the phenomenon, and linguists have theorized in detail about the pragmatic processes involved in implicature recovery and their relation to semantic processes. However, implicature has received less attention from psychologists. Until recently there has been little or no experimental work on implicature, and the topic has not figured prominently in the reasoning literature. This is changing now, however, and the past decade has seen a flurry of experimental work on the psychological basis of implicature, focusing in particular on the questions of whether implicature processing is automatic or effortful and whether pragmatic interpretations develop before or after literal, logical ones. This experimental work is still at a relatively early stage, but some clear findings have emerged, though their significance remains open to interpretation.

In this chapter we review this work, assess its significance, and set it within a wider theoretical context. There are two aspects to this last aim. First, we shall draw attention to a theoretical option largely ignored in the experimental literature. Most experimental work on implicature has been conducted within a broadly Gricean paradigm, according to which implicatures can be calculated and explained using general psycho-social principles. However, there is an alternative, anti-Gricean strand in contemporary philosophy of language, according to which many implicatures, including scalar ones, depend on convention rather than inference. We have no brief for this view, but we think it should not be ruled out and deserves experimental testing. Second, we shall make connections with the literature on dual-process theories of reasoning, as developed by Jonathan Evans and others. Superficially at least, implicature seems made for a dual-process analysis, involving as it does a

* NOTICE: This is the authors' version of a chapter published in K. Manktelow, D. Over, and S. Elqayam (eds), *The Science of Reason: A Festschrift for Jonathan St. B. T. Evans* (pp. 259-81). Psychology Press. It may not reflect changes resulting from the publishing process.

[a] The Open University

[b] The University of Sheffield

contrast of pragmatic and logical responses, and we shall consider if this is correct. We shall close the chapter with some recommendations for future experimental work. Throughout we shall concentrate on scalar implicature, which has been the primary focus for experimenters.

2. Theories of implicature

This section briefly introduces the major theories of implicature. We begin with Grice's account, which provides the context for much subsequent work. Grice held that implicatures are derived by a process of inference, drawing on general principles of human communication, together with the literal meaning of the words uttered, contextual information, and background knowledge. According to Grice, communication is a cooperative enterprise, and hearers assume that speakers will adhere to certain maxims, to the effect that utterances should be: (1) as informative as required (the maxim of quantity), (2) true (the maxim of quality), (3) relevant (the maxim of relevance), and (4) perspicuously phrased (the maxim of manner). Grice proposed that hearers posit implicatures in order to preserve the assumption that speakers are following these maxims. For example, suppose Al asks Bea for her assessment of a student, and Bea replies 'He is very punctual'. This reply violates the maxim of quantity, and Al can preserve the assumption that Bea is being cooperative only by supposing that she wishes to convey something else – that the student is academically weak – which she does not wish to say directly. Al assumes that Bea thought he could work this out, and concludes that she is implicating that the student is weak.

This is an example of what Grice called a *particularized* implicature – that is, it is one that is heavily dependent on the context of utterance (in other contexts Bea's utterance would not generate the implicature mentioned). Grice held that other implicatures were *generalized*; that is, unless explicitly cancelled, the words would generate the same implicature in most contexts. Scalar implicatures are usually regarded as examples of this, and they, too, can be explained in Gricean terms. The speaker chooses a weak term in preference to a stronger one on the same scale – 'some' rather than 'all', for example. But if they had known that the stronger term was applicable, then in not using it they would have been failing to be as informative as required, violating the maxim of quantity. The hearer therefore infers that the speaker did not know that the stronger term was applicable, or, more strongly, that they knew it was not. In this way, Griceans argue, saying that some politicians take bribes can generate the implicature that not all do.

As a cognitive theory, Grice's account has serious limitations. However, it has provided the basis for later theories which aim to give a more cognitively oriented account of implicature derivation. Levinson's work is a particularly detailed example of this so-called *neo-Gricean* tradition (see also e.g. Gazdar, 1979; Horn, 1984, 1989). Levinson focuses on generalized implicatures, which are relatively independent of context. (He acknowledges that particularized implicatures require a different

treatment.) Generalized implicatures, he argues, are *default inferences* – normal or preferred interpretations – which are the product of heuristics, applied at an early stage in language processing, interleaved with the processes of semantic decoding. The inferences are default and go through automatically unless implicitly or explicitly cancelled. This creates an extra level of meaning ('utterance type meaning'), which enriches the content of our utterances in reliable ways, thereby improving the efficiency of human communication.[1]

Levinson identifies three core heuristics, descendants of Grice's maxims. These are:

> The Q-heuristic: What isn't said isn't [the case].
> The I-heuristic: What is expressed simply is stereotypically exemplified.
> The M-heuristic: What's said in an abnormal way isn't normal.
> (Levinson, 2000, pp. 35-8)

We shall concentrate on the Q-heuristic, which is the one that is invoked to explain scalar implicatures. Speakers are presumed to make the strongest statement they legitimately can, with the result that the assertion of a weak claim implies the denial of a salient stronger one. Thus if a speaker chooses a weaker member of a contrast scale in preference to a stronger one, the Q-heuristic generates the inference that the stronger one does not apply. Hence 'Some politicians take bribes' implicates 'Not all politicians take bribes'. The same applies for other scales, such as <*or, and*>, <*possibly, necessarily*>, <*occasionally, often, always*>, and so on. Levinson argues that this view explains a wide variety of linguistic phenomena, including certain facts about lexicalization. For example, we do not have a word meaning *some but not all*, since that meaning attaches by a default inference to the word 'some'.

An alternative approach is *relevance theory* (e.g. Carston, 2002; Sperber and Wilson, 1995; Wilson and Sperber, 2004). Relevance theorists agree with Grice that hearers have expectations of speakers, which motivate the search for implicatures. However, they do away with the apparatus of maxims and heuristics, and replace them with a simple expectation of relevance. On this view, human cognition is automatically geared to maximize the relevance of the inputs it processes, where the relevance of an input is a measure of the positive cognitive effects generated by processing it, set against the effort required to derive them (the 'cognitive principle of relevance'). (A key measure of these positive effects is the number of *contextual implications* an input generates – roughly, how much you can learn from it.) Thus, since speakers want hearers to attend to what they say, utterances carry a presumption of *optimal relevance*, where an optimally relevant stimulus is one that is sufficiently relevant to be worth the hearer's attention and the most relevant the speaker is able

[1] For a related view, see Chierchia (2004), who argues that generalized implicatures are computed by default in the course of grammar processing.

and willing to provide. This presumption (the 'communicative principle of relevance') then guides the hearer's comprehension processes.

Relevance theorists hold that language comprehension is an inferential process; the hearer seeks to infer (rather than decode) the speaker's meaning from their words and the context of utterance. This involves both fleshing out the speaker's literal meaning by resolving ambiguities, identifying references, and so on (a process called 'explicature'), and the search for distinct implicated meanings. The process follows a path of least effort. The hearer forms hypotheses as to what the hearer is seeking to communicate, starting with the most accessible interpretation and progressing to more complex ones, until their expectations of relevance are met or they abandon the search. Implicatures are derived when the literal meaning fails to meet these expectations. For example (and simplifying considerably), suppose Cy asks Dee 'Do all, or at least some, politicians take bribes?' and Dee replies 'Some do'. Since it would clearly be relevant to Cy to know whether all politicians take bribes, in answering thus Dee indicates that she is not in a position to assert that they all do, because (let us assume) she knows it is not true. This further information would increase the relevance of her utterance, so, in line with the principle of relevance, Cy interprets Dee as communicating that some but not all politicians take bribes (see Sperber and Wilson, 1995, pp.277-8).

This example is a scalar implicature. It should be noted, however, that relevance theorists hold that many so-called scalar implicatures are not implicatures at all, properly speaking, but explicatures (e.g. Noveck and Sperber, 2007). In many contexts, they hold, the meaning of 'some' is narrowed down (to, say, *a substantial proportion*) so that it entails *not all*, without generating a distinct *not all* implicature. 'Some' genuinely implicates *not all*, on this view, only where there is an explicit or tacit question as to whether 'all' is applicable, as in the example above. In both cases, however, the same process is at work – a search for an interpretation which meets the hearer's expectations of relevance.

In treating implicatures as derived from psycho-social principles, relevance theory is broadly Gricean (in effect, it assigns all the work to the maxim of relevance). However, there are important differences between it and the neo-Gricean approach, as developed by Levinson. Most importantly, the theories differ on the *order* in which literal and pragmatic interpretations are derived and the processing effort required to derive them. On Levinson's view, the maxims that generate implicatures are applied automatically at an early stage in language comprehension, and pragmatic interpretations are the default ones. Literal interpretations, if they are derived, will require undoing or overriding the default pragmatic inferences. On relevance theory, the opposite is the case. Interpretations are derived in order of accessibility, and in most cases the literal interpretation will be processed first, and a pragmatic interpretation derived only if the literal one fails to meet the hearer's expectations of relevance. A second, related, difference is that relevance theory sees implicature derivation as *context-driven* (Breheny et al., 2006). Expectations of relevance vary with context, and the same utterance may generate an implicature in one context but not in

another, where the hearer's expectations of relevance are different. Thus, from a cognitive perspective, relevance theory makes no distinction between generalized and particularized implicatures; in effect, it treats them all as particularized.

In the experimental literature on implicature the neo-Gricean view and relevance theory are typically the only theories considered. They are certainly the two best worked-out ones from a cognitive perspective. However, there is an important alternative strand in philosophical work on implicature, which rejects the Gricean view that implicatures are derived from the application of general principles. The most detailed development of this anti-Gricean approach is in the work of Wayne Davis (Davis, 1998; see also Morgan, 1978).

Davis treats particularized and generalized implicatures differently. On his view, particularized implicatures (he calls them 'speaker implicatures') exist simply in virtue of the speaker's intentions. For a speaker to mean something beyond the literal meaning of their words is for them to intend to convey that further meaning. Hearers detect particularized implicatures by detecting these intentions, using any of the methods by which mental states can be inferred from behaviour. Generalized implicatures ('sentence implicatures'), on the other hand, depend on semantic *conventions*, according to Davis. It is a convention that sentences of the form 'Some F are G' are used to implicate that not all F are G. Like other conventions, these implicature conventions are to some extent arbitrary. Davis accepts that there will usually be a connection between the literal meaning of a sentence and the implicature it carries which makes the implicature seem fitting. However, he argues that a measure of arbitrariness remains, and that particular implicature conventions cannot be inferred from general principles.

One important argument for Davis's view is that Gricean approaches overgenerate implicatures. For example 'Some politicians take bribes' is a weaker claim than all of the following, which can be thought of as occupying higher points on relevant scales: 'At least 50% of politicians take bribes', 'Some politicians take bribes regularly', 'Some politicians and financiers take bribes' (adapted from Davis, 2008). Yet 'Some politicians take bribes' does not imply the denial of any of those claims. Davis concludes that which denials it implicates is a matter of convention, not inference. Davis also cites evidence that many generalized implicatures are language-specific, drawing on Wierzbicka's work (1985, 1987). For example, in English 'An X is an X' carries the implication that one X is as good as another, but the French translation does not, and the Polish translation implicates that there is something uniquely good about an X (Davis, 1998, p.144, 2008).

Although Davis holds that implicatures are conventional, he does not claim that they are part of the meaning of the sentences that carry them. He distinguishes *first-order* and *second-order* semantic conventions. The former are conventions for the direct expression of thoughts in language; they assign literal meaning to sentences. The latter are conventions for the indirect expression of thoughts by the direct linguistic expression of other thoughts; they assign implicatures to sentences (or rather, to sentence forms; the conventions have some generality). According to Davis,

a language is defined by its first-order rules, not its second-order ones, and implicature conventions are inessential to a language. (He compares them to other conventions of language use, such as that of greeting someone by saying 'How are you?'.) In what follows we shall refer to Davis's view of generalized implicature as *convention theory*.[2]

Note that convention theory cuts across the divisions between the other two approaches. Like relevance theory, it treats the logical meaning of scalar terms as the more basic one. Grasp of logical meanings requires mastery of first-order conventions only, whereas grasp of pragmatic meanings requires mastery of second-order conventions as well. However, like neo-Gricean theory, convention theory treats scalar implicatures as generalized, rather than particularized; once implicature conventions have been mastered they will be applied by default, unless contextually cancelled.

There is a large and complex literature surrounding these theories, drawing on pragmatic intuitions, detailed linguistic analysis, and theoretical considerations about communication and language comprehension. Until recently, however, there was little or no relevant experimental evidence on the topic. As we noted, this is changing, and we expect experimental work to play an increasingly important role in this area. (For useful discussion of the methodology of experimental pragmatics, see Noveck and Sperber, 2007.) The work is still at a relatively early stage, but we feel it is a good time to review the key findings and examine their bearing on the theories mentioned. We turn to this now.

3. Experimental evidence: Development

Experimental work on implicature has focused on scalar implicature, and its main aims have been to establish the reality of the phenomenon and to understand its development. It has been known for some time that children often overlook pragmatic readings of logical terms (e.g. Braine and Rumain, 1981; Smith, 1980). However, it is only recently that specific studies of this effect have been made. The first of these was by Ira Noveck, who argued that children are, in a sense, *more logical than adults* (Noveck, 2001). In one experiment (conducted in French) Noveck tested children's and adults' interpretation of sentences of the form 'Some F are G' where it is known that in fact all F are G (e.g. 'Some giraffes have long necks'). Such statements are *underinformative*, since they assert less than is commonly known to be the case, and they are true on a logical reading but false on a pragmatic one (they are *pragmatically infelicitous*). Noveck used a sentence verification task, presenting participants with samples of underinformative sentences and control sentences and asking whether or

[2] Grice himself recognized the existence of a class of conventional implicatures (1989, pp.25-6). However, he limited this class to those implicatures that are intrinsic to the meaning of the words used, and did not regard the generalized implicatures we are considering as conventional. In Davis's terms, Grice's conventional implicatures are first-order conventions rather than second-order ones (Davis, 1998, p.157).

not they agreed with them. 30 sentences were used, all of the general form 'Some/All Fs are/have/do G'. A sixth were underinformative, the rest were controls that were unambiguously true or false (e.g. 'All chairs tell time', 'Some birds live in cages'). Noveck found that most children accepted the underinformative sentences, even though they correctly evaluated the control sentences. The majority of adults, on the other hand, rejected them. (89% of 8-year-olds and 85% of 10-year-olds accepted the pragmatically infelicitous sentences, whereas only 41% of adults did.) Similar results were obtained with the scalar terms 'Might' and 'Must'. Noveck concluded that a logical interpretation of scalar terms develops before a pragmatic one.

Other studies have replicated Noveck's findings (e.g. Guasti et al., 2005, Experiment 1; Papafragou and Musolino, 2003, Experiment 1; Pouscoulous et al., 2007, Experiment 1). However, the picture is not simple. At least one attempted replication (Feeney et al., 2004, Experiment 1) has failed, with 7-8-year-old children and adults adopting logical interpretations at the same, high, rate (66% and 65%).[3] Moreover, several studies have found that by adjusting the experimental conditions, children's ability to detect implicatures can be much improved. Papafragou and Musolino (2003) made two modifications. First, children were given training to make them more aware of the possibility of pragmatic infelicity (saying 'silly things'), as distinct from outright falsity. For example, they were alerted to the infelicity of describing a dog as 'a little animal with four legs'. Second, the task was adjusted to highlight the informational inadequacy of the underinformative statements. The experiment used stories involving a contest of some kind, where the focus was on the main character's performance – for example, whether he succeeded in putting *all* his hoops around a pole. The result was that the proportion of 5-year-olds rejecting pragmatically infelicitous statements rose dramatically. On some/all tasks, 52.5% rejected the underinformative statements in the modified condition as opposed to only 12.5% in the sentence verification condition. Moreover, like adults, these children justified their answers by invoking statements using the stronger term on the scale ('It was wrong to say some did, since *all* did'). In another study, Papafragou and Tantalou (2004) demonstrated that in naturalistic settings where informational expectations are clear, 4-5-year-olds derive scalar implicatures at high levels (70%-90%). Interestingly, this extended to *particularized* implicatures, dependent on ad hoc, context-dependent scales. For example, a cow was instructed to wrap two gifts, a toy parrot and a doll, which it took out of sight of the participants. On its return, it was asked if it had wrapped the gifts and responded that it had wrapped the parrot. Here the context creates a nonce action scale <*wrapped the parrot*, *wrapped the parrot and the doll*> and the cow's utterance implicates that it had not performed the stronger action. In Papafragou and Tantalou's tests children detected such ad hoc implicatures 90% of the time.

[3] Feeney et al. speculate that this may be due to the fact that their experiment was conducted in English rather than French, and that their results reflect differences in the scope of the French quantifier 'certains' and the English 'some'.

Guasti et al. obtained similar results (Guasti et al., 2005, Experiment 4). They speculated that children might be accepting underinformative statements such as *Some giraffes have long necks*, not because they were failing to draw the scalar inference from *some* to *not all*, but because they were not sure that it was untrue that the statement was false under the pragmatic reading (e.g. that it was untrue that not all giraffes have long necks). To control for this, they used a truth value judgement task (Crain and Thornton, 1998). They asked 7-year-old children to assess descriptions of scenarios (acted by toys and puppets) in which all the relevant information was readily available. For example, they showed children a video in which five soldiers debated whether to ride motorbikes or horses before all deciding to ride horses. They then asked the children whether a puppet was right or wrong to describe the scenario as one in which *some soldiers are riding a horse*. Their main finding was that children's rejection rate for underinformative statements rose virtually to adult levels (75% for children, 83% for adults). (Adult performance on this task also rose, in comparison to that on a simple sentence evaluation task, where only 50% of adults rejected the underinformative statements.) Guasti et al. conclude that 7-year-olds have the ability to derive scalar implicatures, but that it is masked in some experimental settings. In particular, 7-year-olds derive scalar implicatures at adult levels in naturalistic settings where all the relevant information is easily accessible.[4] Guasti et al. note, however, that the same does not go for younger children. 5-year-olds are much less sensitive to scalar implicature than adults, even when tested using a truth value judgement task, with only 50% rejecting underinformative statements (Chierchia et al. 2001; Papafragou and Musolino, 2003). At this age, it seems, at least some children simply lack the cognitive resources to derive implicatures.

Researchers have also begun to investigate other factors that help or hinder implicature derivation in children. Pouscoulous et al. (2007) suggest that implicature derivation is demanding of cognitive resources, and that it can be facilitated by removing distracting factors. To test this, they modified the standard task, by (1) reducing the number of distractors (no unnecessary control questions were used); (2) making the required response an action rather than a metalinguistic judgement (children were asked to adjust the contents of boxes in response to requests from a puppet that all, some, or none should contain items); and (3) using simpler (French) scalar terms ('quelques' was used rather than the more complex 'certains'). The result was an increase in pragmatic responding across the board (at ages 4, 5, 7, and adult), with a developmental progression in performance (logical response rates were 32%, 27%, 17%, and 14% for the four age groups, respectively). In a separate experiment, Pouscoulous et al. confirmed that use of 'quelques' rather than 'certains' increases implicature production in 9-year-olds (logical responses rates were 0% and 42%, respectively).

[4] Guasti et al. also tested the effects of training to avoid pragmatic infelicity. They found that although this did improve children's performance, the effects of the training did not persist beyond the session.

Let us sum up these results. The picture is complex, but the general outline is clear enough: Logical interpretations of scalar terms are more accessible to children than pragmatic ones, and the tendency to adopt pragmatic interpretations (that is, to derive scalar implicatures) increases steadily with age. However, when conditions are right, even relatively young children can derive scalar implicatures, and the process is facilitated in naturalistic settings where informational expectations are clear and all relevant information readily available. Reducing processing demands also helps. The finding that the logical sense of 'some' is developmentally primary is, perhaps, surprising; as Bott and Noveck remark, many people have an intuition that the pragmatic interpretation is more natural (Bott and Noveck, 2004, p.440). Nevertheless, it is well-established and theorists must take account of it. We consider now how these findings bear on the theories of implicature introduced earlier.

First, the data are compatible with relevance theory, and indeed are what the theory would predict. As we noted, on a relevance-theoretic view, interpretations are processed in order of accessibility, starting with the literal meaning, and implicatures are derived only if they are contextually required to satisfy the hearer's expectations of relevance. Thus, on this view, we should expect children to derive implicatures less often than adults, both because they are less sensitive to cues that would raise their expectations of relevance (e.g. about the informational value of possible alternative utterances), and because their cognitive resources are more limited, making implicatures harder to process and so diminishing their relevance (in line with the cognitive principle of relevance). For the same reason, we should expect that pragmatic interpretations would be facilitated by manipulations which serve to highlight the informational gains of alternative interpretations or to reduce computational demands – which is what seems to happen.

The neo-Gricean approach is less easy to reconcile with the data, since it predicts the opposite pattern, with the pragmatic reading of scalar terms being the default one, generated by automatic application of the Q-heuristic. Some explanation will be needed as to why this is not the case for children. Neo-Griceans might appeal to Guasti et al.'s data, arguing that older children at least (7 and upwards) do adopt pragmatic readings by default, and that when they fail to give the corresponding pragmatic *response*, it is because they fail to evaluate the statement correctly under that reading. On this view, the contextual adjustments that facilitate pragmatic responding do so by assisting evaluation rather than derivation. It is unlikely, however, that this explanation can account for all of the effect observed, and it certainly cannot explain the data from younger children. In response neo-Griceans might argue that it takes time for the heuristics to become automatized, and that, until they do, they have to be applied in a slow, effortful fashion. This would also explain why increasing the salience of stronger alternatives and reducing processing demands facilitates inference. We think this neo-Gricean position is a coherent one, though it may be in some tension with Levinson's view that pragmatic processes are closely interleaved with semantic ones (2000, Ch.3). On the view just proposed, the semantic mechanisms would need to be in place first, with the pragmatic ones added later, and

it is not clear how this developmental sequence could result in the interleaving Levinson describes – at least not without temporary disruption to the pre-existing semantic abilities.

What of the third alternative we introduced, the convention theory? This is less easy to assess, given that the theory is not developed from a cognitive perspective. However, prima facie it fits the data well. First, like relevance theory, it correctly predicts the developmental sequence. It will be necessary to master the first-order conventions governing the direct (literal) uses of sentences containing scalar terms before mastering the second-order ones governing their indirect use. (As Davis stresses, second-order conventions are inessential to language proper, and a person could master English without mastering its implicature conventions; Davies, 1998, p.159.) Thus the theory predicts that logical interpretations will precede pragmatic ones developmentally. (We should stress that this prediction is not made by Davis himself, though it follows naturally from his views.) Children who have not mastered the relevant implicature conventions might still be able to derive scalar implicatures, but only by using the methods for recovering particularized, context-dependent implicatures – which, on Davis's view, involve detecting the speaker's intentions. Again, establishing a naturalistic setting with clear conversational goals might be expected to facilitate this process, as the experimental data confirm. Finally, like the neo-Gricean, the convention theorist might argue that children who have mastered the implicature conventions sometimes fail to give appropriate responses because they misevaluate the implicated propositions, appealing again to Guasti et al.'s data.

The convention view also casts new light on data concerning the effect of word choice on implicature derivation. As noted, Pouscoulous et al. (2007) found that French 9-year-olds derive scalar implicatures more readily when 'quelques' is used rather than 'certains', even though both words mean the same (in the sentences used) and should support the same inferences. Pouscoulous et al. explain this effect by arguing that 'certains' is a rarer and semantically more complex word, which soaks up processing resources that could otherwise have been devoted to implicature derivation. However, the effect might also be explained by reference to the second-order conventions associated with the two words. The implicature conventions governing 'certains' might be harder to master than those associated with 'quelques', perhaps because the word is rarer and has more uses. Or the conventions themselves might be more complex. There might even be no settled scalar implicature convention associated with 'certains', with the consequence that any implicatures involving it are particularized. In order to assess such convention-based explanations, systematic study will be needed of the effects of word choice on implicature derivation, both within and across languages.

We conclude that the developmental data reviewed do not conclusively support one theory of implicature over another, though they harmonize better with relevance theory and the convention view. We turn now to other studies, which look at implicature processing in adults.

4. Experimental evidence: Reaction times

A central issue dividing Levinson's neo-Gricean approach and relevance theory is whether logical or pragmatic interpretations are the default ones, and this is something on which reaction time studies should shed light. On Levinson's view logical interpretations require the cancelling or overriding of the default pragmatic interpretations and should therefore, other things being equal, take longer to compute. Relevance theory takes the opposite view and predicts that pragmatic interpretations should take longer.

Bott and Noveck (2004) have carried out important work here, building on earlier findings by Rips obtained in the course of studies of categorization (Rips, 1975). Bott and Noveck ran several experiments to test adult subjects' response times to underinformative sentences. The first experiment used a sentence verification task. Participants saw 54 sentences of the form 'Some/All F are G', half using 'Some', half using 'All'. There were six types of sentence: an underinformative 'some' sentence (e.g. 'Some elephants are mammals'), a straightforwardly true 'some' sentence (e.g. 'Some mammals are elephants'), a straightforwardly false 'some' sentence (e.g. 'Some elephants are insects'), and three 'all' sentences generated by replacing 'Some' with 'All' in the three types of 'some' sentence. Bott and Noveck ran the task twice, on one session instructing participants to treat 'some' logically, as meaning *some and possibly all*, and on the other instructing them to treat it pragmatically, as meaning *some but not all*. They reasoned that if underinformative sentences generate scalar inferences by default, as neo-Griceans claim, then responses to them should take longer on the logical condition than on the pragmatic one, since the initial pragmatic inference would have to be overridden. In fact, the opposite was the case. Correct responses to underinformative sentences took nearly 1400 ms in the pragmatic condition, as opposed to around 800 ms in the logical one. (Responses to control sentences also took longer in the pragmatic condition, but the effect was most marked on the underinformative ones.) Participants also made more mistakes when required to judge underinformative sentences on the pragmatic condition (accuracy rates of 60% as opposed to 90% for the logical condition).

A possible objection to this experiment is that underinformative sentences called for different responses in the logical and pragmatic conditions ('True' and 'False', respectively), and that there might be a response bias in favour of positive answers. To control for this, Bott and Noveck ran a second experiment in which participants were asked to agree or disagree with a second sentence expressing a verdict on the original target sentence ('Mary says the following sentence is true/false…'). The polarity of the verdict was varied with the condition, so that underinformative sentences called for the same response on both conditions. The pattern of results was in line with those from the first experiment.

In a third experiment, Bott and Noveck conducted the test without instructions, allowing participants to adopt whatever interpretation of 'some' they preferred. Again, those who adopted the logical interpretation (classifying underinformative sentences as true) responded more quickly than those who adopted the pragmatic one (2700 ms

and 3300 ms respectively). (This finding was replicated in another study, by Noveck and Posada (2003), where an even more striking difference was found, with pragmatic responders taking nearly twice as long to respond as logical responders.) In a fourth experiment, Bott and Noveck varied the time allowed for responses. In a short-lag condition participants were allowed 900 ms to respond to a presented sentence; in a long-lag condition they were allowed 3000 ms. Rates of logical responses to underinformative sentences fell from 72% in the short-lag condition to 56% in the long-lag one, suggesting that reducing available cognitive resources reduces the likelihood that the scalar inference will be drawn.

These findings suggest that the default reading of 'some' is not *some but not all* and that the derivation of scalar implicatures is effortful and time-consuming, and Bott and Noveck conclude that their studies provide evidence against a default inference account of the kind Levinson proposes. They also note that their experiments provide support for relevance theory – especially the final experiment, which indicates that the likelihood of deriving a scalar implicature varies with the availability of cognitive resources, as relevance theory predicts.

We do not deny that these findings offer support for relevance theory, but we wish to strike a note of caution. We have four points to make. First, Bott and Noveck's data may look different when viewed through the lens of dual-process theory. It may be that different tasks activate different *types* of reasoning process, and this needs to be taken into account in interpreting the resulting response times. We shall discuss dual-process theories in the next section.

Second, at most, the reaction-time data show that pragmatic readings are not derived *before* logical interpretations. They do not show that they are not derived by the application of the Q-heuristic. The heuristic might be applied *after* a logical interpretation has been derived and its application might be relatively effortful. Of course, the claim that the pragmatic reading is the initial one is central to Levinson's account, but a weaker neo-Gricean position might modify or drop it. Such an account might continue to hold that neo-Gricean heuristics are applied *by* default, in the sense that they are applied routinely unless contextually cancelled, even though the pragmatic reading is not itself the initial ('default') one.

Third, the data are broadly compatible with convention theory, as well as with relevance theory. Like relevance theory, convention theory holds that the logical reading of scalar terms is the more basic one, and it predicts that the derivation of pragmatic interpretations will be more effortful. Deriving a logical reading involves applying first-order conventions only, whereas deriving a pragmatic one involves applying both first-order and second-order conventions. Of course, without some account of the cognitive processes involved, it is difficult to make predictions from this, but, prima facie, one would expect the latter process to take longer. Like the neo-Gricean, the convention theorist might still hold that implicature conventions, once mastered, are applied by default, unless contextually cancelled.

Fourth, there is a possible alternative explanation for the delayed response times to underinformative sentences in the pragmatic condition, which does not attribute it to

the cost of deriving the pragmatic interpretation. The response time to a sentence is the sum of two components: the time taken to derive the preferred interpretation and the time taken to evaluate it, and (as Bott and Noveck themselves acknowledge), it is possible that the delay in responding to underinformative sentences on the pragmatic condition is attributable to the latter component rather than the former. In these cases, the inferred proposition is of the form *Some but not all F are G*, where the hearer in fact knows that all F are G (e.g. *Some but not all elephants are mammals*). If evaluating this proposition involves a search for a (non-existent) counterexample (e.g. an elephant that is not a mammal), then this may well be the source of much of the delay in responding. (It might explain the high error rate, too.) On the logical condition, by contrast, the proposition to be evaluated is of the form *At least one F is G* (e.g. *At least one elephant is a mammal*), which is easily verified by finding a single confirming instance.

There is some support for this explanation. One interesting feature of the data from Experiment 1 is that response times for underinformative *some* statements on the pragmatic condition were longer, not only than corresponding responses on the logical condition, but also (by around 400 ms) than responses to control *some* statements *on the same, pragmatic, condition*. It is implausible to attribute this delay to the time taken to derive the pragmatic interpretation, since participants were under instructions to adopt that interpretation for all sentences on this condition. It is more likely that it is due to the additional time required to evaluate the proposition derived in the underinformative cases. Bott and Noveck assume that the scalar inference is unique to underinformative (pragmatically infelicitous) statements – the 'T1' statements in their materials (see e.g. 2004, p.451). We do not understand this, however. Why wouldn't the control *some* statements (e.g. 'Some mammals are elephants', 'Some elephants are insects') prompt the inference as well? At any rate, they should on neo-Gricean or convention-based accounts, and it would be begging the question against those accounts to suppose otherwise. Moreover, it would be absurd to suggest that in real life situations, people only use 'some' to imply *not all* if they know that the *not all* statement is false![5]

It may be objected that this explanation cannot account for the data from Bott and Noveck's fourth experiment, where time pressure produced a majority of logical responses to underinformative sentences. If participants derive the pragmatic interpretation in the short-lag condition but lack time to evaluate it, then we should expect their responses to be at chance, rather than showing a bias towards the logical answer. We acknowledge the power of this consideration but still feel that caution is necessary. For example, it is possible that under pressure participants defaulted to

[5] We wonder if the term 'underinformative' fosters some confusion. On a Gricean view, scalar implicatures arise when a less informative term is chosen in preference to a more informative one. In a sense, then, scalar implicatures arise from underinformative statements. But this sense of 'underinformative' is different from the one used in the experimental literature, where a *some* statement is said to be underinformative if the corresponding *all* statement is commonly known to be true. Underinformativeness in this sense is not a prerequisite for scalar implicature on any account.

processing an association (elephant → mammal) or that the pressure induced a bias towards positive answers, which was not present in the other conditions (the logical response here was 'True'). We conclude that although Bott and Noveck's data offer some support for relevance theory, the other theoretical options remain live, in particular convention theory.[6]

5. A dual-process perspective

The studies reviewed so far have been conducted largely in isolation from work on reasoning, and few connections have been made with the large body of work on 'dual-process' theory – a field in which Jonathan Evans has been a pioneer. This work may be highly relevant, however, and we turn to it now. We begin with a brief introduction to dual-process theories. We should stress that this is highly simplified. In particular, many dual-process theorists are now recognizing that the neat binary divisions that have been proposed are too crude and require refinement and qualification (see the papers in Evans and Frankish, 2009). For present purposes, however, the following characterization will be sufficient.

There is abundant evidence for the existence of two separate but interacting types of processing in human reasoning, decision making, and social cognition, which may deliver different and sometimes conflicting results (for reviews, see Evans, 2003, 2008; Frankish and Evans, 2009). One type of processing (referred to as 'implicit', 'tacit', 'heuristic', 'experiential', or simply 'type 1') is fast, effortless, automatic, nonconscious, inflexible, and contextualized. The other ('explicit', 'analytic', 'type 2') is slow, effortful, controlled, conscious, flexible, and decontextualized. Type 1 processes are often characterized as parallel and as either associative or based on 'fast and frugal' heuristics, whereas type 2 processes are usually described as serial and rule-based. In the field of reasoning, dual-process theories were originally devised to explain evidence from deductive reasoning tasks, where subjects' responses often reveal a conflict between logical processes and non-logical biases. The paradigm case of this is *belief bias*: a non-logical preference for believable over unbelievable conclusions, which interferes with the goal of selecting valid over invalid conclusions (Evans et al., 1983). In the dual-process framework, belief bias is regarded as a type 1 process, whereas logical responses are ascribed to type 2 processing. Thus, our capacity for type 2 reasoning is seen as the source of our ability to decontextualize problems and respond in accordance with logical norms.

Some researchers have gone on to develop dual-*system* theories of mental architecture, which integrate work on reasoning, decision-making, and learning.

---

[6] Another group of reaction-time studies, by Breheny et al. (2006), which used more naturalistic materials, escapes our third objection at least. In one experiment, Breheny et al. presented scalar sentences at the end of short vignettes, one creating a *lower-bound* context, in which the scalar implicature was not appropriate, the other an *upper-bound* context, in which it was. They found that participants took longer to read the scalar sentences in the upper-bound contexts than in the lower-bound ones, suggesting that scalar implicatures are not generated by default, but only where contextually needed.

These theories ascribe the two types of processing to distinct cognitive systems, usually known as *System 1* and *System 2*, which have different functions, processing styles, and evolutionary origins (Evans and Over, 1996; Sloman 1996, 2002; Stanovich 1999, 2004). One important claim is that individual differences in general intelligence and working memory are differences in the capacity of System 2, not System 1 – a claim defended at length in Stanovich's work (1999, 2004; Stanovich and West, 1998, 2001). For example, Stanovich and West (1998) showed that the ability to solve decontextualized versions of the Wason selection task, which require analytic processing, correlates with high IQ, whereas performance on heavily contextualized versions, which can be solved by nonconscious heuristics, is unrelated to IQ.

How does implicature processing fit into this framework? Since implicatures are pragmatic interpretations, a natural view would be to see them as the product of type 1 processes, and to regard logical interpretations as the result of type 2 reasoning – at least in adults, in whom such reasoning is well developed. There is some reason for thinking this may be the case. Although adults are sensitive to scalar implicatures, it is notable that, on all studies, they make some logical responses as well – sometimes at a high rate. They might be simply failing to derive the implicature, of course, perhaps because their expectations of relevance are met by the logical reading. But dual-process theory offers an alternative explanation: that they are deriving the implicature but inhibiting the response it dictates and responding in line with a logical interpretation generated by type 2 processes. This is not implausible. It would not be surprising if underinformative scalar statements often provoke explicit, type 2 reasoning, at least when presented without a natural conversational background. For the implicatures these statements generate are by definition known to be false, and in conversation one would use such statements only when one wanted to, as it were, *lie by implicature*. As Guasti et al. note (2005, pp.690-2), this may cause some adult participants to depart from standard conversational norms, and either infer that the experimenter is using scalar terms in a technical, strictly logical sense, or search for counterexamples that would render the implicature true (in the latter case, the 'logical' response would in fact manifest a pragmatic reading). Although Guasti et al. do not make a connection with dual-process theories, this process might well involve explicit, type 2 reasoning.

The hypothesis that logical adult responders are inhibiting pragmatic responses has been tested by Feeney and collaborators (Feeney et al. 2004). Using a computerized sentence verification task, they measured the responses and reaction times of 50 adults to underinformative statements and controls (the materials used were similar to those in Noveck's original experiments). They found that logical responses to infelicitous (i.e. underinformative) *some* statements took significantly longer than logical responses to felicitous ones (i.e. ones that are true under a pragmatic reading, such as 'Some cars are red'). This suggests that the former responses are preceded by additional processing, which, Feeney et al. propose, involves the drawing and inhibiting of a pragmatic inference. Support for this, they argue, comes from a secondary finding that the tendency to respond logically to

infelicitous *some* statements is positively correlated with cognitive capacity, as measured by a counting-span task – though they note that this result should be treated with caution owing to the small sample size. Again, this suggests that extra processing is involved in generating the logical responses, and Feeney et al. make an explicit link with dual-process theory. Feeney et al. also propose that their findings offer support for neo-Gricean theory, since they suggest that it is the undoing rather than the derivation of a scalar implicature that is effortful.[7]

We find these results suggestive but also puzzling. As noted earlier, we fail to see why felicitous *some* statements would not also prompt scalar implicatures, and if they do, then the time difference in logical responses to felicitous and infelicitous *some* statements cannot be attributed to the presence or absence of a pragmatic inference. However, this is not incompatible with Feeney et al.'s suggestion that the difference reflects the costs involved in inhibiting a pragmatic response. For in the infelicitous case, the pragmatic and logical interpretations dictate different responses, and if a subject switches from a pragmatic to a logical reading they will need to inhibit their initial disposition to answer 'False'. In the felicitous case, by contrast, the pragmatic interpretation dictates the same response ('True') as the logical one. Thus, even if a participant makes the same switch of readings, they should still respond more quickly, since they will not need to inhibit their initial, pragmatically driven, disposition to answer 'True'. (And, of course, if they do not make the switch, their response should be quicker still.) We feel this explanation – which is consistent with Feeney at al.'s overall conclusion – is the more plausible one.

In another study, Scrafton and Feeney (2006) used a dual-process framework to investigate the development of scalar implicature. They point out that there is evidence that type 1 processes develop before type 2 ones, with the former being well developed by age 10, but the latter continuing to mature up to age 15 (Handley et al., 2004; Klaczynski 2001).[8] If scalar implicatures are generated by type 1 processes, and logical responses (in adults) by type 2 ones, then, Scrafton and Feeney argue, certain age-related effects should be evident. First, in children, when type 1 processes are still developing, implicature detection should be patchy but facilitated by contextual cues. Second, in young adolescents pragmatic responding should be dominant regardless of context, since type 1 processes are then well-developed but type 2 ones still maturing. Third, logical responding should re-emerge in adulthood, as type 2 processes become fully developed.

Scrafton and Feeney tested these predictions by comparing sensitivity to scalar implicature among five groups, of age 6, 9, 12, 15, and adult, using both contextually impoverished materials (sentences based on those used by Noveck (2001)) and

---

[7] Feeney et al.'s results were notable for the high rates of logical responding to underinformative *some* statements found among adults (77%, with half the participants giving logical responses only). We speculate that the abstract nature of the sentence verification task fostered an analytic approach in participants.

[8] Scrafton and Feeney use the terms 'heuristic' and 'analytic' rather than 'type 1' and 'type 2', but we shall continue to use the latter, which are less loaded.

contextually enriched materials (storyboards in which, for example, a girl was shown eating all the sweets on the table and then telling her mother that she had eaten *some* of them). The results confirmed the predictions. 6-year-olds detected implicatures only in the enriched condition. 9-year-olds detected implicatures in both conditions, but detected more in the enriched one. 12-year-olds detected implicatures in both conditions at near ceiling. Finally, among 15-year-olds and adults the rate of logical responses increased again and context had little effect, suggesting that the logical responders were decontextualizing the task. These developmental data thus fit well with a dual-process framework. Of course, as Scrafton and Feeney note, it is highly unlikely that young children's logical responses are due to type 2 processes. When they respond logically, it is likely that they are doing so because they have not yet developed the type 1 processes needed to derive the pragmatic reading, rather than because they possess the type 2 processes needed to inhibit it.

How does this dual-process framework fit with the theories of implicature reviewed earlier? We think it is broadly compatible with all of them. The framework involves no commitment as to the nature of the type 1 processes involved in implicature derivation, except that they should fit the general type 1 profile – fast, nonconscious, effortless, etc. They might involve application of neo-Gricean heuristics, relevance-based processes, or second-order conventions. It might be objected that relevance theory should be excluded, since it treats implicature derivation as effortful and logical interpretation as effortless. However, it is important to distinguish two senses of 'effortful'. One is simply 'computationally costly', and in this sense we can certainly speak of type 1 processes as being more or less effortful. In the dual-process literature, by contrast, 'effortful' is used to refer to processes that load on working memory, and in this sense effortful reasoning is, virtually by definition, a type 2 process. Now, we assume that in discussions of relevance theory the term is used in the first sense. Relevance theorists think of pragmatic processes as spontaneous and nonconscious, rather than conscious and reflective (e.g. Sperber and Wilson, 1995, Ch.2; Wilson and Sperber, 2004). Thus we do not see a conflict here between dual-process theory and relevance theory. The same goes for relevance theory's claim that logical interpretations are effortless. We need to distinguish between those logical interpretations that result from a failure to derive an implicature and those that result from inhibiting one. It is the former that relevance theory treats as effortless (first sense), and the latter that dual-process theory treats as effortful (second sense).

However, although we think that relevance theory is formally compatible with dual-process theory, there is one aspect of Scrafton and Feeney's work that does offer selective support for neo-Gricean theory or convention theory. This is the finding that by early adolescence context ceases to have any effect on implicature derivation. That is, implicature derivation moves from being a context-driven process to being a largely default one, albeit one that can be overridden. This finding is hard to reconcile with relevance theory, according to which implicature derivation is always context-

driven, and it harmonizes better with neo-Gricean or convention-based accounts, on which the process involves mastering general principles or conventions.

So far, we have assumed that implicature derivation is a type 1 process, but can it also be a type 2 process (effortful in the second sense)? De Neys and Schaeken have argued that it can (2007). They used a dual-task methodology to see if cognitive load on working memory interfered with pragmatic processing. Subjects undertook a sentence verification task while performing a dot-pattern memorization task, once with a simple control pattern, once with a complex load pattern. De Neys and Schaeken found that there was a modest though significant decrease in the rate of pragmatic interpretations under the complex load (76% vs. 70%), but no decrease in the rate of correct responses to the control sentences. They also found that pragmatic responses took significantly longer under load (by about 700 ms), whereas responses to control sentences were not affected. De Neys and Schaeken conclude that their findings contradict the neo-Gricean account and support relevance theory, indicating that implicature derivation is effortful, not automatic, and that people are more logical under cognitive load. They also conclude that, since the pattern memorization task loaded on working memory (that is, was effortful in the second sense), implicature derivation also draws on working memory.

We have three points to make here. First, the fact that a task draws on working memory is not sufficient to show that it involves type 2 reasoning. There are many processes which involve attention, and hence working memory, but which do not involve explicit, type 2 thought processes (see Barrett et al., 2004, for a long list of processes associated with working memory capacity).[9] And it is plausible that additional attentional resources are needed for responding pragmatically to underinformative sentences (for one thing, as noted earlier, their evaluation is not simple). Second, even if type 2 reasoning were involved in implicature derivation, this would not fit well with relevance theory. On a relevance-theoretic view, language comprehension is a complex process of non-demonstrative inference, involving the parallel forming and testing of hypotheses about explicit content, implicated premises, and implicated conclusions, and drawing on expectations about specific cognitive effects as well as a general assumption of relevance. This process is not one that could feasibly be performed by a slow, conscious, decontextualized, serial reasoning system. If implicature derivation were a type 2 process, it is more likely that it would involve the application of simple heuristics. Thus De Neys and Schaeken's conclusion, were it sound, would actually favour a form of neo-Gricean view, albeit one which allows that heuristics can be applied effortfully as well as automatically. Third, although we think it implausible that implicature derivation is typically a type 2 process, we do not exclude the possibility that it may sometimes be. In particular, utterances that are hard to interpret would be likely to provoke explicit, conscious reasoning, perhaps involving the application of general conversational principles such as those Grice

[9] Thanks to Jonathan Evans for this point.

proposed, or beliefs about the rules governing scalar terms.[10] Effortful (second sense) implicature derivation may occur in some experimental settings, too. For example, in Bott and Noveck's (2004) Experiment 1, where participants were instructed to interpret 'some' pragmatically, it is likely that executing this instruction would involve type 2 activity. (Responsiveness to verbal instruction is often regarded as a distinctive mark of type 2 processing; e.g. Evans and Over, 1996.) It is important to recognize this possibility, and to control for it.

Dual-process approaches to implicature are still in their infancy, and it is too soon to draw firm conclusions. The findings reviewed are tantalizing but need replication, and they are also partially at odds with some of the work discussed earlier. In particular, the claim that adults' logical readings of underinformative sentences are often the product of type 2 reasoning fits ill with the data from reaction-time studies. Here we can merely highlight the need for further work. Dual-task methodologies and searches for correlations between performance and measures of working memory capacity should be useful. It will also be important to consider the nonconscious control processes responsible for initiating type 2 reasoning and resolving conflicts between the two systems – processes which Jonathan Evans has dubbed *type 3* (Evans, 2009).

Conclusions

Our conclusions are modest. The experimental work reviewed establishes the psychological reality of scalar implicature and is highly relevant to the evaluation of competing theories of implicature processing. However, the data are not wholly consistent, and they are compatible with all three major accounts. There is some support for relevance theory over the neo-Gricean approach, particularly from reaction-time studies, but it is not decisive. The current experimental findings are also consistent with convention theory, which has an attractive economy and fits the developmental data well, combining the view that logical interpretations are basic with the view that pragmatic interpretations follow a developmental progression from context-driven to default. We shall conclude with some suggestions emerging from our discussion.

First, psychologists should give consideration to convention theory, developing cognitive models of convention-based processes and subjecting them to experimental testing. Studies of the effect of word choice on implicature derivation should be particularly useful here, since the theory predicts that implicatures will be specific to particular linguistic constructions. Cross-linguistic data should also be highly relevant. Second, it is important to devise methods of assessing the relative contribution of

[10] De Neys and Schaeken suggest that standard dual-process theory cannot allow that implicature derivation is a type 2 process, since it attributes pragmatic biases to the automatic system and normatively correct responses to the effortful one. However, it is now widely accepted that type 2 reasoning can generate non-normative responses, and that it can involve other procedures besides the application of logical rules. See the papers in Evans and Frankish, 2009.

different types of process to the generation of responses to scalar statements, including deviation processes, evaluation processes, and inhibitory processes. Different computational paths may lead to similar overt responses and reaction times, and it is important to find ways of distinguishing them. Third, there is need for caution in the experimental use of underinformative statements. Though handy as touchstones, such statements are anomalous, in that the implicatures they generate are always false, and this may distort responses and reaction times. It is also important to remember that felicitous *some* statements generate scalar implicatures, too. Finally, we believe that work on scalar implicature will benefit from closer integration with the reasoning literature and in particular with dual-process theories. There is evidence that type 2 reasoning is responsible for logical interpretations of scalar terms in adults, and it could play a role in generating pragmatic interpretations, too. In moving to a dual-process framework, theories of implicature will need to become more complex, allowing for a plurality of processes and methods, and taking account of the factors which trigger them.

Acknowledgements

The authors wish to thank Shira Elqayam and Jonathan Evans for helpful discussions about implicature and dual processes.

References

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, *130*, 553-573.

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.

Braine, M., & Rumain, B. (1981). Children's comprehension of "or": Evidence for a sequence of competencies. *Journal of Experimental Child Psychology*, *31*, 46-70.

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

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

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and beyond: The cartography of syntactic structures, Volume 3* (pp. 39-103). New York: Oxford University Press.

Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A., & Meroni, L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In A. H. Do, L. Domínguez, & A. Johansen (Eds.), *Proceedings of the 25th annual Boston University conference on language development* (pp. 157-168). Sommerville, MA: Cascadilla Press.

Crain, S., & Thornton, R. (1998). *Investigations in universal grammar*. Cambridge, Mass.: MIT Press.

Davis, W. A. (1998). *Implicature: Intention, convention, and principle in the failure of Gricean theory*. Cambridge: Cambridge University Press.

Davis, W. A. (2008). Implicature. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy (Winter 2008 edition)*. Retrieved April 7, 2009, from http://plato.stanford.edu/archives/win2008/entries/implicature/.

De Neys, W., & Schaeken, W. (2007). When people are more logical under cognitive load: Dual task impact on scalar implicature. *Experimental Psychology*, *54*, 128-133.

Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*, 454-459.

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255-278.

Evans, J. St. B. T. (2009). How many dual-process theories do we need? One, two, or many? In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 33-54). Oxford: Oxford University Press.

Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition*, *11*, 295-306.

Evans, J. St. B. T., & Frankish, K. (Eds.) (2009). *In two minds: Dual processes and beyond*. Oxford: Oxford University Press.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove: Psychology Press.

Feeney, A., Scrafton, S., Duckworth, A., & Handley, S. J. (2004). The story of some: Everyday pragmatic inference by children and adults. *Canadian Journal of Experimental Psychology*, *58*, 121-132.

Frankish, K., & Evans, J. St. B. T. (2009). The duality of mind: An historical perspective. In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 1-29). Oxford: Oxford University Press.

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

Grice, P. (1989). *Studies in the way of words*. Cambridge, Mass: Harvard University Press.

Guasti, M. T., Chierchia, G., Crain, S., Foppolo, F., Gualmini, A., & Meroni, L. (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes*, *20*, 667-696.

Handley, S. J., Capon, A., Beveridge, M., Dennis, I., & Evans, J. St. B. T. (2004). Working memory, inhibitory control and the development of children's reasoning. *Thinking and Reasoning*, *10*, 175-195.

Horn, L. R. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (Ed.), *Meaning, form, and use in context* (pp. 11-42). Washington, DC: Georgetown University Press.

Horn, L. R. (1989). *A Natural history of negation*. Chicago: University of Chicago Press.

Klaczynski, P. A. (2001). Analytic and heuristic processing influences on adolescent reasoning and decision-making. *Child Development*, 72, 844-861.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, Mass.: MIT Press.

Morgan, J. L. (1978). Two types of convention in indirect speech acts. In P. Cole (Ed.), *Syntax and semantics, Volume 9: Pragmatics* (pp. 261-280). New York: Academic Press.

Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*, 165-188.

Noveck, I. A., & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language*, *85*, 203-210.

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

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

Papafragou, A., & Tantalou, N. (2004). Children's computation of implicatures. *Language Acquisition*, *12*, 71-82.

Pouscoulous, N., Noveck, I. A., Politzer, G., & Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition*, *14*, 347-375.

Rips, L. J. (1975). Quantification and semantic memory. *Cognitive Psychology*, *7*, 307-40.

Scrafton, S., & Feeney, A. (2006). Dual processes, development, and scalar implicature. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th annual conference of the Cognitive Science Society* (pp. 774-779). Austin, TX: Cognitive Science Society.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3-22.

Sloman, S. A. (2002). Two systems of reasoning. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 379-396). Cambridge: Cambridge University Press.

Smith, C. L. (1980). Quantifiers and question answering in young children. *Journal of Experimental Child Psychology*, *30*, 191-205.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd edition). Oxford: Blackwell.

Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. Lawrence Erlbaum Associates: Mahwah, NJ.

Stanovich, K. E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin*. Chicago: University of Chicago Press.

Stanovich, K. E., & West, R. F. (1998). Cognitive ability and variation in selection task performance. *Thinking and Reasoning*, *4*, 193-230.

Stanovich, K. E., & West, R. F. (2001). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*, 645-665.

Wierzbicka, A. (1985). Different cultures, different languages, different speech acts. *Journal of Pragmatics*, *9*, 145-178.

Wierzbicka, A. (1987). Boys will be boys: "Radical semantics" vs. "radical pragmatics". *Language*, *63*, 95-114.

Wilson, D., & Sperber, D. (2004). Relevance theory. In L. R. Horn & G. L. Ward (Eds.), *The handbook of pragmatics* (pp. 607-632). Oxford: Blackwell.