Effortful Pragmatics: The Demanding Nature of Implicatures

Wim De Neys (Wim.Deneys@psy.kuleuven.be)
Walter Schaeken (Walter.Schaeken@psy.kuleuven.be)
Department of Psychology, K.U.Leuven, Tiensestraat 102
B-3000 Leuven, Belgium

Abstract

In the present study we introduce dual task methodology to test opposing psychological processing predictions concerning the nature of implicatures in pragmatic theories. Implicatures routinely arise in human communication when hearers interpret utterances pragmatically and go beyond the logical meaning of the terms. The neo-Gricean view (e.g., Levinson, 2000) assumes that implicatures are generated automatically whereas Relevance Theory (Sperber & Wilson, 1996) assumes that implicatures are effortful and not automatic. Participants were presented a sentence verification task with underinformative sentences that have the potential to produce scalar implicatures like *Some oaks are trees*. Depending on the nature of the interpretation of *Some* (logical or pragmatic) the sentence is judged true or false. In two experiments, cognitive working memory resources were experimentally burdened by the concurrent memorization of complex dot patterns during the interpretation process. Results showed that participants made more logical and fewer pragmatic interpretations under load. Findings provide direct support for the Relevance Theory view.

Introduction

The complexity of human communication has been characterized as one of the hallmarks of our species. A striking demonstration of the sophisticated nature of our communication system is our ability to draw pragmatic inferences or implicatures. Often speakers intend to convey far more than the logical meaning of the words they utter and hearers readily retrieve the intended interpretation. Imagine, for example, that U.S. Secretary of Defense Donald Rumsfeld states in an interview:

(a) Some of our soldiers are misbehaving.

You would readily assume that Mr. Rumsfeld intends to say:

(b) Not all of our soldiers are misbehaving.

Rather then:

(c) All of our soldiers are misbehaving.

However, according to standard semantic accounts, *Some* means *Some and possibly All*. From a strictly logical point of view, Secretary Rumsfeld's utterance is therefore compatible with (c) and not with (b). In statement (b), *Some* is interpreted pragmatically. Based on expectations about how a conversational exchange occurs the hearer goes beyond the literal, logical meaning of the uttered terms (Grice, 1989). Such pragmatic inferences or implicatures have been extensively studied in the philosophical and linguistic work on verbal communication. The inference in the example has been dubbed *scalar implicature* since the constituting terms can be ordered on a scale of informativeness. The implicature arises when a less informative term (e.g., Some) is taken to imply the denial of the more informative term (e.g., All). The scalar implicature is considered as the paradigmatic case for the study of implicatures.

In the literature there is an interesting discordance between pragmatic theorist who try to describe how scalar inferences actually work (i.e., how a scalar inference manifests itself in real time). The dispute centers around the automaticity of the scalar implicature (see Bott & Noveck, 2004; Noveck, 2001; Noveck & Posada, 2003; Papafragou & Musolino, 2003). One school of thought, referred to as the "neo-Gricean" account (e.g., Levinson, 2000) claims that the pragmatic interpretation is actually the default interpretation in a concrete, communicative setting. It is assumed that every time *Some* is encountered the hearer will automatically make the scalar implicature. Levinson, for example, has argued that scalar implicatures result from a Q-heuristic dictating that "What isn't said isn't". Therefore, whenever a listener hears a weak scalar term like *Some*, the listener will automatically assume that the speaker intended that a stronger term (i.e., All) is not warranted. The Q-heuristic dictates that if it is not said that *All* is the case, it is not. If the speaker intended *Some* to imply the stronger term *All*, she should have said it explicitly. Hence, the hearer will always start by making the scalar implicature. An eventual logical interpretation of *Some* can only arise in a later stage where the implicature is undone by the context.

A second approach, Relevance Theory (Carston, 1999; Sperber & Wilson, 1996), assumes implicatures are not made by default. Relevance theory states that how far a hearer goes in processing an utterance's meaning is governed by principles concerning effect and effort. Listeners try to gain as many effects as possible for the least effort. According to Relevance Theory the logical

interpretation of *Some* (i.e., some and possibly all) could very well lead to a satisfying interpretation in an utterance. It is possible that the hearer will derive the scalar implicature and move to a pragmatic interpretation (i.e., some but not all) to make the utterance more informative but this enrichment is not automatic and will come at the cost of additional processing effort.

Thus, from the perspective of Relevance Theory scalar implicatures are considered as effortful, non-necessary inferences, whereas according to a neo-Gricean view it is precisely the occasional undoing of an automatic, default implicature that takes extra effort (Noveck & Posada, 2003).

Intuitively, the neo-Gricean account seems to be the most plausible. If one looks at Mr. Rumsfeld's utterance (a), for example, it certainly *feels* like we infer (b) directly while the logical interpretation (c) seems to require far more active thinking. However, contrary to the intuitive appeal, recent experimental findings seem to provide some support for the Relevance Theory view. Noveck (2001), for example, showed that children made the logical interpretation of *Some* more than adults (see also Papafragou & Musolino, 2003). Thus, the availability of a larger cognitive resource pool with increasing age seemed to be linked with a stronger dominance of the pragmatic interpretation. In a series of latency studies, Bott and Noveck (2004, see also Noveck & Posada, 2003) observed that people needed more time for the pragmatic interpretation than for the logical one. They also found that limiting the time available for responding boosted the rate of logical interpretations. These findings suggest that making scalar implicatures is associated with correlates of effortful processing.

The present study introduces a dual task approach to settle the debate with a more direct test of the automaticity claim. In two experiments available cognitive resources were experimentally burdened by imposing a resource demanding secondary task during sentence interpretation. Participants were presented so-called underinformative sentences like *some tuna are fish* and were asked to judge whether the sentence was true or false (e.g., Bott & Noveck, 2004). Since, by definition, all tuna are fish, the sentence will be judged false if one draws the scalar implicature and interprets the sentence pragmatically (e.g., some but not all tuna are fish). If one interprets *Some* logically (e.g., some and possible all tuna are fish) the sentence will be judged true.

If the scalar implicature is not automatic but requires effortful processing, making the implicature should be harder when cognitive resources are burdened. Therefore, from the perspective of Relevance Theory one predicts that the rate of pragmatic inferences will decrease under cognitive load. On the other hand, if one believes that the implicature is made automatically and it is the logical interpretation that requires additional processing, burdening the cognitive resources should hamper the logical interpretation process. Hence, people should be more likely to stick to the default pragmatic interpretation under load. From a neo-Gricean view one therefore predicts that the cognitive load will boost the rate of pragmatic interpretations.

Notice that pragmatic theorists and previous experimental studies have not characterized the exact nature of the alleged effortful processing. The present study focuses on the role of executive working memory resources since these are widely recognized as the quintessential component of human cognitive capacity (e.g., Engle, Tuholski, Laughlin, & Conway, 1999).

Experiment 1

In Experiment 1 participants verified sentences while they concurrently tried to remember a briefly presented visual dot pattern (e.g., Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). The complexity of the dot pattern was manipulated so that storage of the pattern in a control group would be less demanding.

Method

Participants

A total of 184 students of the University of Leuven, Department of Social Sciences, participated voluntarily or in return for psychology course credit.

Material

Sentence verification task. Participants provided True/False judgments for 10 underinformative sentences presented on a computer screen in front of them. Content of the sentences referred to categories and exemplars (e.g., Some <*exemplar*> are <*category*>). The only instructions participants were given was to respond 'True' if they thought the sentence on the screen was true, or 'False' if they believed the sentence to be false. Participants typed down their response using the numpad ('1'- True or '2'- False).

Participants also judged 10 filler trials where patently true (e.g., Some birds are eagles) and patently false (e.g., Some tigers are fish) sentences were presented (see Appendix for the complete set). The sentences were presented in random order.

Dot memory task. The dot memory task is a classic spatial storage task (e.g., Bethell-Fox & Shepard, 1988; Miyake et al., 2001). For the present study a 3x3 matrix filled with three to four dots was briefly presented for 850 ms. Participants memorized the pattern and were asked to reproduce it afterwards.

In the load group the matrix was filled with a complex four-dot pattern (i.e., a "two- or three-piece" pattern based on the work of Bethell-Fox & Shepard, 1988, and Verschueren, Schaeken, & d'Ydewalle, 2004, see Figure

1). Miyake et al. (2001) established that storage of similar complex dot patterns tapped executive resources.

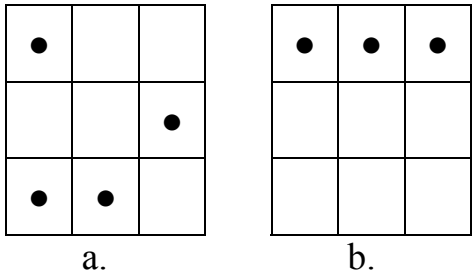

Figure 1. Examples of the dot patterns in the load (a) and control group (b).

In the control group the pattern consisted of three dots on a horizontal line (i.e., a "one-piece" pattern in Bethell-Fox & Shepard's terms). This simple and systematic pattern should only minimally burden the executive resources (e.g., De Neys, in press; Miyake et al., 2001).

Procedure

Participants were tested in groups of 19 to 42 and were randomly assigned to the control or complex load group. The experiment started with a demonstration of the storage task. On two practice storage items (one with a simple and one with a complex pattern) an empty response matrix was presented 1 s after the pattern had been presented. Participants used the keypad to indicate the location of the dots. Instructions stressed that it was crucial that the dot patterns were reproduced correctly in the upcoming task.

A verification task trial started with a brief presentation of the dot pattern for 850 ms. Next, the sentence was presented and remained on the screen until the participant made a response. Afterwards, an empty matrix was presented and participants had to reproduce the dot pattern. Participants received feedback on whether the pattern had been reproduced correctly and were reminded that they had to remember the complete pattern correctly. The procedure was clarified with two practice sentences.

Results and discussion

Dot memory task. Results for the dot memory task indicated that the task was properly performed. The mean number of correctly localized dots for the complex four-dot pattern in the load group was 3.74 (SD = .27) and 2.88 (SD = .18) for the simple three-dot pattern in the control group. Thus, overall, about 94% of a complex and 96% of a simple pattern was reproduced correctly. This establishes that participants were indeed memorizing the patterns.

Sentence verification task. Underinformative sentences. Remember that when underinformative sentences are interpreted pragmatically they will be judged "false" in the sentence verification task. Overall, participants gave a mean number of 7.76 "false" responses to the underinformative sentences. Thus, the vast majority (about 78%) of participants' interpretations were pragmatic in nature. The crucial finding is that the rate of pragmatic interpretations tended to be lower in the complex load (M = 7.33, SD = 3.64) group than in the control group (M = 8.18, SD = 3.25), $t(182) = 1.65$, $p < .1$, $d = .24$. Although the effect reached only marginal significance the trend is consistent with the prediction from a Relevance Theory perspective. Concurrent memorization of the complex dot pattern decreases rather than boosts the tendency to make pragmatic interpretations.

Filler sentences. To check whether the load did not interfere with basic sentence comprehension we calculated the mean number of correct responses to the filler sentences. Performance was uniformly high, both in the control (Mean = 9.32, SD = .73) and load group (Mean = 9.18, SD = .86), $t(182) = 1.25$, $p = .25$. The high overall rate of correct filler responses (less than 7% errors) establishes that participants were not simply guessing or superficially accepting under secondary task load. In principle, one could suggest that the secondary task load triggers a general tendency to select "true" responses. Hence, the load effects on the underinformative sentences would not point to a difficulty to derive the implicature *per se*. Note, however, that four out of ten filler sentences were patently true. Hence, if participants would simply accept or guess under load they should err on 50% to 60% of the filler trials. The high rate of correct responses on the filler items establishes that the cognitive burden specifically interfered with drawing scalar implicatures.

Experiment 2

The marginal load trend observed in Experiment 1 presented some initial support for the Relevance Theory view of implicatures. Experiment 2 presents a more stringent within-subjects test to validate the findings. Participants verified two sets of sentences, one while memorizing the easy control patterns and one while memorizing the complex load patterns.

Method

Participants

A total of 56 first-year psychology students of the University of Leuven participated in return for psychology course credit. None of the students had participated in Experiment 1.

Material

Sentence verification task. Participants verified two sets of 20 sentences (with 10 underinformative and 10 filler sentences in each set, see Appendix).

Dot memory task. See Experiment 1.

Procedure

Participants were tested in small groups. For each participant it was randomly determined with which sentence set and which load type they started. Item presentation followed the same procedure as in Experiment 1.

Results and discussion

Dot memory task. As in Experiment 1, participants properly memorized the dot patterns. The mean number of correctly localized dots for the complex four-dot pattern was 3.75 (SD = .25) and 2.90 (SD = .17) for the simple three-dot pattern. Thus, overall, about 94% of a complex and 97% of a simple pattern was reproduced correctly.

Sentence verification task. Underinformative sentences. Overall, the vast majority (about 76%) of participants' interpretations were pragmatic in nature. However, as in Experiment 1, participants made significantly fewer pragmatic interpretations when they had to memorize the demanding complex patterns (M = 7.32, SD = 3.98) than when memorizing the easy control patterns (M = 7.89, SD = 3.68), $t(55) = 2.18$, $p < .04$, $d = .41$.

Filler sentences. Participants responded correctly to the filler sentences both when memorizing the control (Mean = 9.36, SD = .77) and complex patterns (Mean = 9.3, SD = .1.01), $t(55) = .37$, $p = .72$. Hence, all results completely replicated the findings of Experiment 1.

General Discussion

The present study introduced dual task methodology to test opposing processing assumptions in pragmatic theories concerning the nature of scalar implicature. The intuitively appealing neo-Gricean account considers implicatures as automatic, default inferences that will need to be effortfully overridden to arrive at a logical interpretation. From the perspective of Relevance Theory the logical interpretation is more basic and implicatures are considered additional, non-automatic, cognitive demanding inferences. The present findings clearly supported the Relevance Theory view of scalar implicature. Both in Experiment 1 and 2, burdening the cognitive resources with the memorization of the complex dot patterns decreased the rate of pragmatic inferences. Hence, contrary to intuition, people became more "logical" under cognitive load. If it would be the logical interpretation that required additional effortful processing, one would have expected to see more pragmatic and fewer logical interpretations under complex load. The decreased pragmatic interpretations directly establish that making scalar implicatures is not automatic but requires effortful, cognitive processing.

One might note that the impact of our cognitive load manipulation was not spectacular. Indeed, the complex dot memorization only decreased the rate of pragmatic interpretations with about 10% and participants still made over 70% pragmatic interpretations under load. However, it should be clear that the size of the load effect is not the crucial issue here. First, one could always try to obtain stronger effects by imposing more demanding secondary tasks (e.g., increasing the number of dots). Second, and more important, it is obvious that no one in the Relevance Theory camp would claim that drawing an implicature puts a *massive* burden on our cognitive resources. Given the prevalence of pragmatic inferences in daily life, it would indeed be hard to see how people would manage to communicate if every implicature would involve a major cognitive cost. Thus, it is rather likely that implicatures require a relatively small cognitive involvement. The important point is that they are not completely automatic. The basic finding that a cognitive burden results in more logical and fewer pragmatic interpretations directly contradicts the prediction from a neo-Gricean account of scalar implicature.

We stated that pragmatic theorists have not yet characterized the exact nature of the alleged effortful, cognitive processing. The memorization of complex dot-patterns has been shown to specifically burden the executive component of the human working memory system (Miyake et al., 2001). The present findings thus indicate that the effortful processing that is required in deriving scalar implicatures specifically draws on these executive working memory resources. This stipulation should stimulate pragmatic theorists to link their work more closely with the rich psychological research tradition on the role of executive resources in higher-order cognition (e.g., Kane & Engle, 2002). We believe that such an approach will be especially fruitful to arrive at more fine-grained future characterizations of the cognitive operations underlying pragmatic inferences.

We finally note an interesting implication of the present findings for the field of deductive reasoning. It is well-established that fallacious deductive inferences can be often attributed to pragmatic interpretations of the premises (e.g., Begg & Harris, 1982; Brain & O'Brien, 1998; Johnson-Laird & Byrne, 2002). Influential dual process theories of reasoning (e.g., Evans, 2003; Stanovich & West, 2000) have attributed this pragmatic modulation to the operation of a heuristic system that biases the operation of a second, analytic system mediating normative correct reasoning. The basic assumption is that the heuristic system operates

automatically whereas the second system would be cognitively demanding and draw on executive working memory resources. The present study indicates that contrary to the popular conceptualization, pragmatic modulation is not a pure automatic process. People already need executive resources to derive the potentially biasing scalar implicatures. Consistent with recent findings in the reasoning field (e.g., De Neys, Schaeken, & d'Ydewalle, 2005) this implies that the basic characterization of an automatically operating heuristic-pragmatic reasoning system should be questioned. Our findings clearly established that the paradigmatic case for the study of pragmatics, the scalar implicature, is not made automatically but involves effortful, executive processing.

Acknowledgements

Preparation of this manuscript was supported by a post doctoral fellowship grant from the Fund for Scientific Research-Flanders (FWO).

References

Begg, I., & Harris, G. (1982). On the interpretation of syllogisms. *Journal of Verbal Learning and Verbal Behavior, 21,* 595-620.

Bethell-Fox, C. E., & Shepard, R. N. (1988). Mental rotation: Effects of stimulus complexity and familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 12-23.

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. D. S., & O'Brien, D. P. (Eds.). (1998). *Mental logic*. Mahwah, NJ: Erlbaum.

Carston, R. (1999). Informativeness, relevance, and scalar implicature. In R. Carston, & S. Uchida (Eds.), *Relevance theory: Applications and implications*. Amsterdam: John Benjamins.

De Neys, W. (in press). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science.*

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2005). Working memory and everyday conditional reasoning: Retrieval and inhibition of stored counterexamples. *Thinking & Reasoning, 11,* 349-381.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*, 309-331.

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

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

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review, 109*, 646-678.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual differences perspective. *Psychonomic Bulletin & Review, 9*, 637-671.

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

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General, 130,* 621-640.

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.

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

Sperber, D., & Wilson, D. (1996). *Relevance: Communication and cognition*. Oxford: Basil Blackwell.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences, 23*, 645-726.

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2004). Everyday conditional reasoning with working memory preload. *Proceedings of the Twenty-Sixth Annual Meeting of the Cognitive Science Society*. Mahwah: Erlbaum.

Appendix

Sentences Used in the Verification Tasks

<u>Set 1 (Experiment 1 and 2)</u>

Underinformative

Some eels are fish.
Some carp are fish.
Some oaks are trees.
Some beeches are trees.
Some sparrows are birds.
Some robins are birds.
Some flies are insects.
Some mosquitoes are insects.
Some roses are flowers.
Some tulips are flowers.

Filler

Some birds are magpies. (true)
Some insects are wasps. (true)
Some pigeons are insects. (false)
Some beetles are flowers. (false)
All Chrysanthemum are flowers. (true)
All hazels are trees. (true)
All trees are elms. (false)
All fish are herrings. (false)
All daffodils are trees. (false)
All sycamores are fish. (false)

Set 2 (Experiment 2)

Underinformative

Some ants are insects.
Some bees are insects.
Some canaries are birds.
Some blackbirds are birds.
Some daisies are flowers.
Some lilies are flowers.
Some firs are trees.
Some birches are trees.
Some trout are fish.
Some sharks are fish.

Filler

Some flowers are carnations. (true)
Some trees are willows. (true)
Some crocuses are trees. (false)
Some poplars are fish. (false)
All cod are fish. (true)
All parrots are birds. (true)
All birds are crows. (false)
All insects are worms. (false)
All pike are birds. (false)
All swallows are insects. (false)