Processing pragmatic inferences

Bob van Tiel

1 Introduction

One assumption that permeates much of the literature is that pragmatic inferences can be categorised into a number of well-defined classes. Each of these classes is associated with a set of criteria that a proposition has to fulfill to be categorised as a bona fide member of that class: presuppositions must project, conversational implicatures must be cancellable, calculable, and nondetachable, conventional implicatures must be lexicalised and truth-conditionally inert, et cetera.

Insofar as this assumption is on the right track, one might wonder whether these theoretical distinctions have some kind of correspondence in behavioural data. In other words, one might wonder whether different kinds of pragmatic inferences are associated with differences in response patterns and reaction times. In order to evaluate this *correspondence hypothesis*, I compared eight types of pragmatic inferences within a verification task. This contribution is based on previous work I did in collaboration with Walter Schaeken (van Tiel & Schaeken 2015).

2 Eight kinds of pragmatic inferences

In this verification task, participants read sentences that were followed by pictures. It was their task to indicate whether the sentence was an appropriate description of the subsequent picture. In the target condition, the sentence was followed by a picture that falsified an inference that might have been triggered by the sentence. I assume that participants who compute the inference judge the sentence 'false' and participants who do not judge the sentence 'true'. In addition, each sentence type was presented in control situations in which it was unambiguously true or false. Figure 1 shows for each inference type an example sentence, its hypothesised pragmatic inference, and a target situation in which the sentence is true on its unenriched interpretation but in which the hypothesised inference is false.

In total, we tested eight types of pragmatic inferences. First, three types of scalar inferences; one based on the ⟨or, and⟩ scale and two based on the ⟨some, all⟩ scale. Second, three types of inferences that are often explained as varieties of quantity implicature: distributivity inferences, conditional perfection, and exhaustivity in 'it'-clefts. A more extensive discussion of these inferences can be found in van Tiel & Schaeken (2015). Third, the uniqueness presupposition associated with definite

Pre-proceedings of 'Trends in Experimental Pragmatics', XPRAG.de, Berlin, Germany, pp. 146–152.

descriptions. Lastly, a syntactically ambiguous sentence for which the stronger surface scope reading entails the weaker inverse scope reading.

If the correspondence hypothesis is correct, it is expected that the theoretical boundaries between these four groups of inference types will be visible in truth judgements, verification times, or perhaps even both.

3 The experiment

3.1 Participants

40 students at the Université Libre de Bruxelles, all native speakers of French, participated in the experiment for financial compensation (33 females, mean age: 21, range: 18-28).

3.2 Materials

The experiment consisted of 62 items and included eight types of sentences corresponding to the eight types of pragmatic inferences. For each inference type, three kinds of situations were constructed: two control situations and one target situation. In the first control situation, the sentence was unambiguously true; in the second control situation it was unambiguously false; in the target situation its truth value depended on whether the pragmatic inference was derived. See Figure 1 for examples of target situations.

Target situations occurred four times for each inference type; control situations twice. The order of the items was randomized for each participant.[1]

3.3 Procedure

On each trial, the target sentence was displayed first. Participants were instructed to press the space bar as soon as they had read and understood the sentence. Thereupon, the sentence disappeared and was replaced by a picture. Participants had to decide as quickly as possible whether the sentence was true or false as a description of the depicted situation, and had to register their decision by pressing one of two keys. Thereupon, the picture disappeared and was replaced by the message '(Press the space bar to continue.)'. Upon pressing the space bar, the next trial commenced.

1 I included four other types of pragmatic inferences. For various reasons, these had to be excluded from the analysis.

Scalar inference 'some'
Some of the shapes are red.
⇝ Not all of the shapes are red.

Scalar inference 'some not'
Some of the shapes are not red.
⇝ Some of the shapes are red.

Scalar inference 'or'
The square or the circle is red.
⇝ Not both of them are red.

Distributivity inference
Each of the shapes is red or green.
⇝ Not all of the shapes are red.

Conditional perfection
Each shape is red if it is a circle.
⇝ Not all of the shapes are red.

Exhaustivity in 'it'-clefts
It is the circle that is red.
⇝ Only the circle is red.

Uniqueness presupposition
The circle is not red.
⇝ There is only one circle.

Syntactic ambiguity
All of the shapes are not red.
⇝ None of the shapes are red.

Figure 1 Target sentences followed by their pragmatic inferences and a situation that falsifies the inference but verifies the unenriched interpretation of the target sentence.

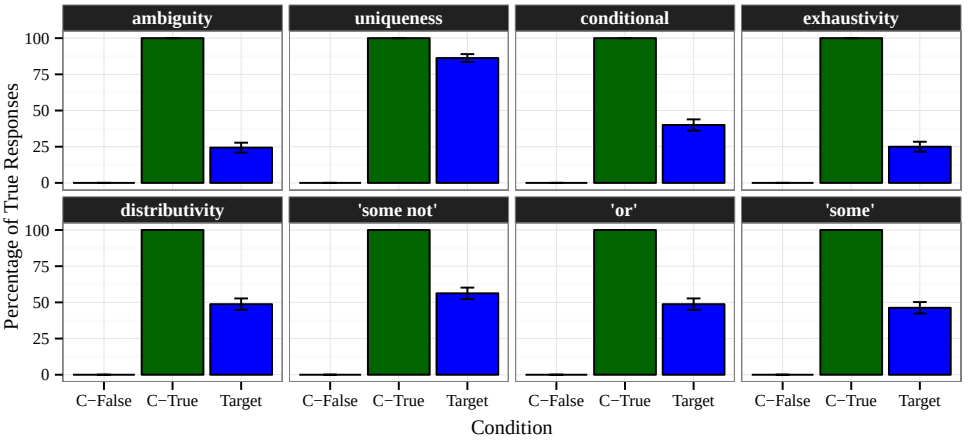

Figure 2 Percentages of 'true' responses for each type of pragmatic inference. Error bars represent standard errors.

3.4 Data treatment

All participants and trials were included in the analysis. Decision times were logarithmised to reduce the skewness of their distribution.

3.5 Choice proportions

The percentages of 'true' responses for each inference type are summarised in Figure 2. For scalar inferences (46% 'true' responses for 'some', 49% for 'or', and 56% for 'some not'), conditional perfection (40%), and distributivity inferences (49%) 'true' responses were roughly as frequent as 'false' responses. Participants in the case of exhaustivity in 'it'-clefts (25%) and syntactic ambiguities (24%) had a pronounced preference for 'false' responses. Conversely, in the case of uniqueness presuppositions (86%), there was a pronounced preference for 'true' responses.

In general, error rates were quite low, with one exception: in the case of 'some not', many participants judged the sentence false in a situation in which it was true on both its literal and enriched reading. I do not have a straightforward explanation for this anomaly. Perhaps participants considered the sentence too unspecific to be considered an appropriate description of the situation.

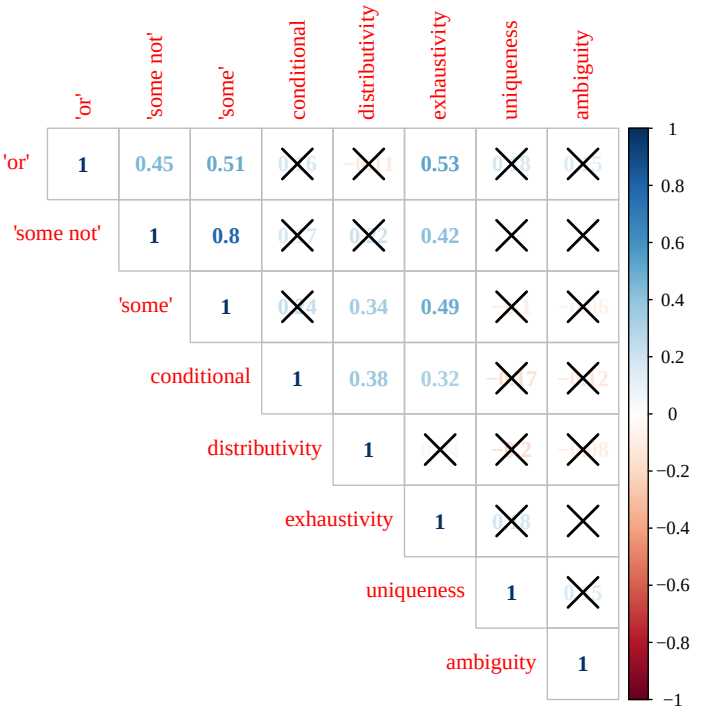

Figure 3 Product-moment correlation coefficients for each pair of inference types. Correlations for which $p > .05$ are crossed out.

3.6 Correlations

One might wonder whether participants were consistent across different inference types, i.e. whether participants gave a comparable number of 'false' responses for each inference type. To test this, we calculated the product-moment correlation between the number of 'false' responses for each pair of inference types. The results of this analysis are summarised in Figure 3.

There were significant correlations between each pair of scalar inferences, between exhaustivity in 'it'-clefts and each kind of scalar inferences, between exhaustivity in 'it'-clefts and conditional perfection, between conditional perfection and distributivity inferences, and between distributivity inference and scalar inferences associated with 'some'. None of the other correlation coefficients were significant.

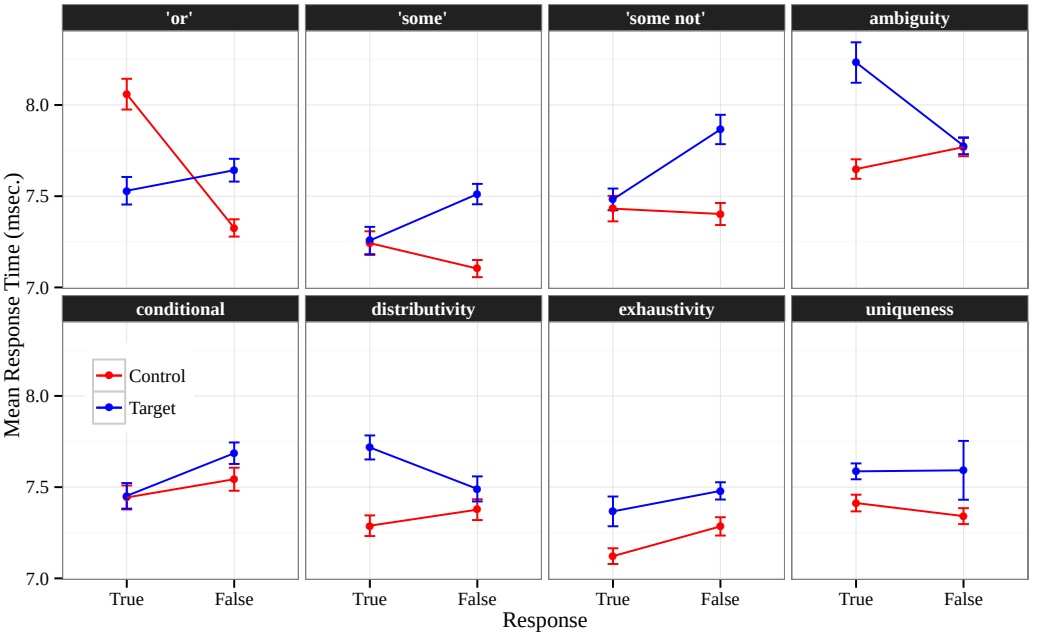

Figure 4 Mean logarithmised decision times for each type of pragmatic inference. Error bars represent standard errors.

3.7 Decision times

Decision times refer to the time from the onset of the picture to the button press indicating the truth judgement. For the analysis of the decision times, all error trials were removed. Mean logarithmised decision times are summarised in Figure 4.

Mixed models were constructed predicting logarithmised decision times based on condition (target or control) and response ('true' or 'false'), including random intercepts for participants and items. Degrees of freedom and corresponding p-values were calculated based on Satterthwaite's procedure as implemented in the `lmertest` package. I assume that an interaction between condition and response indicates that the computation of a pragmatic inference facilitated or delayed decision times.

The interaction between condition and response was significant for each kind of scalar inference: 'or' (β = -0.77, SE = 0.26, t = -3.00, p = .022), 'some' (β = -0.34, SE = 0.11, t = -3.00, p = 0.03), and 'some not' (β = -0.45, SE = 0.12, t = -3.6, $p < .001$). It was also significant, but in the opposite direction, for distributivity inferences (β = 0.31, SE = 0.11, t = 2.76, p = .006) and ambiguous sentences (β

= 0.45, SE = 0.11, t = 4.09, $p < .001$). It was not significant for any of the other inference types (all p's $> .2$).

In summary, the computation of scalar inferences was associated with a processing cost. Calculating distributivity inferences and arriving at the stronger surface scope reading of the ambiguous sentence facilitated decision times. There were no effects on decision times for the computation of conditional perfection, exhaustivity in 'it'-clefts, and the uniqueness presupposition of definite descriptions.

4 Discussion

The correspondence hypothesis predicts that theoretical distinctions between different types of pragmatic inferences are reflected in experimental data, in casu the results of a verification task. This hypothesis was partly confirmed. Scalar inferences behaved relatively homogenously: participants were highly consistent across different kinds of scalar inferences and their computation was associated with a delay in decision times, unlike other kinds of pragmatic inferences.

To some extent, the three varieties of quantity implicature also patterned together, as can be seen in the correlation plot in Figure 3. However, this pattern was clearly less strong and was not associated with a particular processing signature. The uniqueness presupposition and syntactically ambiguous sentence behaved differently from the three types of quantity implicature in terms of choice proportions and consistency but not in terms of verification times.

Hence, although experimental data can provide an insight into the provenance of a type of pragmatic inference, there does not seem to be a specific litmus test to arrive at a decisive verdict. Rather, data from different measures should be taken into consideration. In order to learn more about what measures are relevant for which types of pragmatic inference, however, it will be necessary to extend the scope to different kinds of presuppositions and ambiguities.

References

van Tiel, Bob & Walter Schaeken. 2015. Processing conversational implicatures: alternatives and counterfactual reasoning. To appear in: *Cognitive Science*.