



Focus Influences the Presence of Conditional Perfection: Experimental Evidence

Marie-Christine Farr
Goethe University Frankfurt
Marie.C.Farr@gmail.com

Abstract. This contribution tests whether focus has a determining influence on the occurrence of Conditional Perfection. Two off-line experiments support this hypothesis. Conditional Perfection occurs significantly more often if the antecedent of the conditional is focused compared to the non-focused case. Additionally, in contrast to the scalar implicature associated with *or* (Zondervan 2009), Conditional Perfection occurs only infrequent if the antecedent is not focused. The second experiment suggests that this distinct behavior is due to different properties of the scalar implicature and the implicature associated with Conditional Perfection.

1 Introduction

Conditional Perfection (CP) describes the phenomenon that speakers interpret conditional sentences, under certain conditions, as biconditionals. The phenomenon was given its name by Geis & Zwicky (1971), who observed that their students extended or “perfected” the meaning of conditionals. Consider Geis & Zwicky’s original example:

- (1) a. If you mow the lawn, I’ll give you five dollars.
- b. If you don’t mow the lawn, I won’t give you five dollars.
- c. Only if you mow the lawn, I’ll give you five dollars.

As McCawley (1993) points out, “Only if p, q” can be paraphrased as “If not p, not q”. So examples (1b) and (1c) should be regarded equivalents. CP thus described the inference from (1a) to (1b) or (1c). Geis & Zwicky note that perfection of conditionals is clearly wrong from a logical viewpoint. The utterance of a sentence like (1a) does not exclude the possibility that the hearer gets a reward for some other effort, for example, for cleaning the living room. Still they affirm that many speakers interpret conditionals in exactly this way and that CP is “highly regular” (Geis & Zwicky 1971: 564). This paper addresses the conditions that influence the occurrence of CP.

2 Conditional Perfection Is a Pragmatic Phenomenon

Standard approaches analyze conditionals within the framework of possible worlds semantics, as in (2):

- (2) “If p , q ” is true in w if and only if:
 $\forall w' \text{ such that } w' \in C(w) \ \& \ p(w') : q(w')$
 with w being the actual world and $C(w)$ being the set of possible worlds accessible from w .

A conditional “If p , q ” is considered to be true in w , if for all possible worlds that are accessible from w and in which p is true q is also true. No information is incorporated in the conditional meaning about what will be the case if p does not hold. So if we want to keep the semantics for conditionals unchanged, we must assume that CP is pragmatic. To affirm this claim, conditionals that do not allow perfection offer convincing support. If CP was contingent upon the semantics it must arise with all conditionals. *Biscuit Conditionals*, as the standard example taken from Horn (2000), belong to this group:

- (3) a. If you’re thirsty, there’s some beer in the fridge.
 b. If you’re not thirsty, there’s no beer in the fridge.

The conditional in (3a) clearly does not invite the inferences in (3b). The reason for this is that the two propositions are conditionally unrelated. Further evidence in favor of a pragmatic analysis of CP provides the fact that the inference is usually cancelable. That is, it can be “taken back” by adding additional information. The following example illustrates this (van Canegem-Ardijns & van Belle 2008):

- (4) If you mow the lawn, I’ll give you five dollars.
 But also if you paint the garage.

Through adding the additional condition under which five dollars will be paid, the CP inference that mowing the lawn is the only possible way to earn the money is canceled.¹ Another feature suggesting a pragmatic account of CP is expressed by Boër & Lycan (1973). They support the idea that not all conditionals, in all situations are perfected and give the following counterexample:

- (5) If John quits, he will be replaced.

¹ As discussed in van Canegem-Ardijns & van Belle (2008: 371 ff.), this cancels just the “Only if p , q ” inference. Canceling the “If not p , not q ” inference is often not possible for speech acts like promises, threats or warnings. This might indicate that the two inferences “If not p , not q ” and “Only if p , q ” are indeed not equivalent. Still, we will stick to this assumption.

Here, according to Boër & Lycan, CP does not arise, because common sense inhibits the implication that *if John does not quit his job, he won't be replaced*. It is easily imaginable that John will be replaced, e.g. because his work lacks quality or because he does not get along with his colleagues. This constitutes the main argument of Boër & Lycan. CP does not arise because of the special form or intrinsic features of conditionals but simply because of additional information like world or common-sense knowledge.

All this provides convincing evidence that CP is a pragmatic rather than a semantic inference.

3 Conditional Perfection Is not a Scalar Implicature

Originally, Geis & Zwicky claimed that it is difficult to explain CP in terms of implicatures (cf. Grice 1989) and that the inference is clearly not a conversational implicature. They argue that Grice, when characterizing conversational implicatures

looks for general principles governing the effects that utterances have, principles associated with the nature of the speech itself. CP is, in some sense, a principle governing the effects that utterances have (...) but it is in no way that we can see derivable from considerations having to do with the nature of the speech act. (Geis & Zwicky 1971: 565)

Later in their squib, Geis & Zwicky relativize this claim and state that an explanation of CP in terms of conversational implicatures is not *easily* establishable. Unfortunately they conclude without further elaboration or clarification. Despite this rejection, it is the most popular assumption today to explain CP with respect to conversational implicatures. One particularly favored approach is to analyze CP in terms of *Scalar Implicatures* (SI). The general idea of a SI is that an inference can be drawn based on the amount of information that is expressed. As Grice's first maxim of Quantity advises, speakers should make their contribution as informative as required. Thus, very simplified, if we know that more informative statements than the one actually made exist (and some further assumptions hold (cf. e.g. Geurts 2010)) we can argue that the speaker does not believe that the more informative statements are true. For this reason, *Horn Scales* are created to order terms according to their information content. In the following, three scales will be introduced that were suggested to account for CP.

3.1 Attempt 1: Atlas & Levinson (1981)

Atlas & Levinson (1981) discuss the following scale:

if and only if p, q
 \uparrow *if p, q*

There are two problems with this scale: First and most important, this scale is inadequate to explain CP. As mentioned above, by uttering a weaker statement the speaker expresses that he does not believe that the stronger statement holds. Thus an effect opposing CP arises. Uttering a conditional would always imply that *p* is only one among more conditions which will bring about *q* (Horn 2000). Atlas & Levinson assumed that this scale does not elicit SIs, hence they did not see this problem. They argued that the scale is not well-formed because the two terms are not lexicalized to the same degree. This is why the unwanted SI does not arise. However, this restriction does not hold, as will be discussed below. The second deficiency of the above scale is pointed out by van der Auwera (1997,b). The element at the top of the scale is just the literal meaning plus the SI which is expected. It is clear that this combination will always be more informative than the literal meaning on its own. Since a construction of this type is excluded with respect to other scalar terms like <some, some but not all>, it should also be excluded for CP. A related weakness is the complexity of the statement at the top of the scale. A restriction which is often proposed for potential alternatives is that they must not be considerably longer or more complex than the statement made. In these cases, the speaker could just choose the shorter statement in order to be brief. However, this restriction does indeed not hold. As Matsumoto (1995) points out, more informative statements need to be asserted if they contain relevant information and even if they are of a greater complexity. Therefore, the above scale does elicit the unwanted SI. However, the scale can be ruled out due to other reasons such as the above-mentioned deficiencies and additional constraints on monotonicity.²

3.2 Attempt 2: van der Auwera (1997b)

Van der Auwera (1997; 1997b) assumes the following scale for his approach:

...
if p, q and if r, q and if s, q
if p, q and if r, q
 \uparrow *if p, q*

If someone utters “If *p, q*” the comparably stronger statements such as, e.g., “If *p, q* and If *r, q*” are automatically denied and hence the speaker expresses that *p*

² Horn scales must not include items of different monotonicity behavior. (cf. e.g. Matsumoto 1995).

is the only condition that will bring about q . The weakness of van der Auwera's scale is that he has to assume an infinitely large expression at the very top of the scale that incorporates all possible antecedents. Only if the statement "If p , q " negates *for all* other antecedents r that "If r , q " is true, CP arises. The problem of this account is the reference to particular antecedents. Someone who hears a conditional like (1a) would have to reason for all possible antecedents that they will not bring about q . This however, does not seem appropriate to represent the reasoning involved in CP. It is implausible and probably impossible that someone who hears a conditional has an infinitely long list of possible antecedents in mind (Horn 2000; von Fintel 2001).

3.3 Attempt 3: Horn (2000)

Horn (2000) suggests yet another approach. He believes that the CP effect is due to pragmatic strengthening and suggests the following scale:

$$\begin{array}{c} q / \textit{Whatever the case, } q \\ \uparrow \textit{if } p, q \end{array}$$

Unfortunately, a proposal like Horn's is also not sufficient to derive CP, as von Fintel (2001) notes. Uttering "If p , q ", and thereby negating the simple statement $q/\textit{whatever the case, } q$, does not mean that p is the only necessary and sufficient condition under which q will occur. It only elicits the much weaker implicature that q is not unconditionally true. Whether only one or several conditions exist that render q true cannot be decided at that point. Thus Horn's account is missing an important step towards CP. Still, following Horn, CP is derivable under his account (Horn 2000, 2004). Horn performs a reduction of Grice's maxims which is motivated by the desire to be in accordance with the idea of a dualistic functional model which guides conversation. This model assumes that utterances are subject to two forces. All of Grice's maxims and submaxims, except for the maxim of Quality which remains unchanged, can be reduced to express just these two forces. Horn ends up with the following two principles (Horn 1993):

- **Q Principle:** Make your contribution sufficient. Say as much as you can, given quality and the R Principle.
- **R Principle:** Make your contribution necessary. Say no more than you must, given the Q Principle.

The **Q Principle** embodies the first maxim of Quantity and the first two submaxims of Manner. It corresponds to a lower bound on information content. Within the **R Principle** the maxim of Relation, the second maxim of Quantity and the third and fourth submaxims of Manner are collected. This principle

constitutes an upper limit on the form of the utterance.

Under these prerequisites uttering “If p , q ” implies that the speaker does not believe that q is unconditionally true. Given the **R** Principle, this contribution must be more relevant to the conversation than uttering q alone. This is why CP arises. As Horn puts it: “what could make such a condition [i.e. “If p ”] more relevant than its necessity?” (Horn 2000: 310). Thus, CP occurs naturally and regularly due to systematic strengthening which is evoked by the **R** Principle. A serious problem of Horn’s account is that it would never be relevant to claim that p is just a sufficient condition. This is however, wrong. With *mention-some* questions it is typically sufficient to provide only the most relevant out of several conditions (cf. von Stechow 2001).

The attempts to explain CP as a SI have not been satisfactory. The next section discusses the possibility to analyze CP still as a *Quantity Implicature* (QI) but not as a SI.

4 Conditional Perfection as Quantity Implicature

The basic idea, following von Stechow (2001), is that CP is a QI³ which arises as a by-product of an exhaustive interpretation. The following examples, taken from von Stechow (2001) and Groenendijk & Stokhof (1984) illustrate exhaustive interpretations.

- (6) a. Q: Who left the party early?
A: Robin and Hilary left the party early.
b. Robin and Hilary but no one else left the party early.
- (7) a. Q: Who walks?
A: John and Mary walk.
b. John and Mary but no one else walk.

If the answers in (6a) and (7a) are interpreted exhaustively, they correspond to (6b) and (7b). These inferences are thought to be QIs. However, they cannot be derived by the “standard procedure” introduced for SIs, i.e. by negating stronger statements. The problem is the set of relevant alternatives. In order to infer “John and Mary but no one else walk” from “John and Mary walk” an infinitely large set of more informative statements needed to be rejected, as illustrated below:

- John, Mary and Peter walk.

³ Although QIs entail SIs, I use QI in the following to refer to all kinds of QIs except for SIs.

- John, Mary, Ann and Peter walk.
- ...

This is again very implausible.

Van Rooij & Schulz (2004) and Schulz & van Rooij (2006) provide an account of exhaustive interpretation which overcomes these problems. Their theory of exhaustification captures the intuition that exhaustive reasoning is based on the *closed world assumption*. Details of their theory are not important for our purpose; what is important is that they explain exhaustification with the use of minimal models. An exhaustive interpretation corresponds to a minimal model. Consider the following models for example (7a).

M_1	M_2
Individuals = {Robin, Ben, John, Mary}	Individuals = {Robin, Ben, John, Mary}
$\llbracket walk \rrbracket^{M_1} = \{\text{Robin}\}$	$\llbracket walk \rrbracket^{M_2} = \{\text{Ben}\}$

M_3	M_4
Individuals = {Robin, Ben, John, Mary}	Individuals = {Robin, Ben, John, Mary}
$\llbracket walk \rrbracket^{M_3} = \{\text{John, Mary}\}$	$\llbracket walk \rrbracket^{M_4} = \{\text{John, Mary, Ben}\}$

Table 1: The table presents four potential models for example (7a)

Exhaustification arises with a model that makes the answer true and in which the extension of the predicate in question is minimal. From the above given examples M_3 is the model leading to an exhaustive interpretation. The others either make the answer false (M_1 and M_2) or the extension of *walk* is not minimal (M_4). An important feature of this account is that the focus-background partitioning determines the predicate in question that gets minimalized. Therefore, minimalization and hence exhaustification applies only to focused terms. In this regard Rooth (1996) makes an interesting observation with respect to focus and the nature of question-answer pairs. When we consider question-answer pairs, the position of focus in the answer corresponds to the *wh*-constituent in the question. Examples (8a) and (8b) illustrate this.⁴

⁴ Boldness equals focus in this and further examples.

- (8) a. Q: Who walks?
 A: **John and Mary** walk.
 b. Q: What do John and Mary do?
 A: John and Mary **walk**.

Due to the different focus-background partitioning, different models will be minimal. For (8a) a model in which only John and Mary walk is minimal, but for (8b) a model where John and Mary do nothing else but walk is minimal.

4.1 Predictions for Conditional Perfection

Based on what we have seen so far, CP can be explained in terms of exhaustification which in turn can be explained by the selection of a minimal model. Since exhaustification is sensitive to focus, the same conditional can be interpreted exhaustively and non-exhaustively dependent on the question of the contexts and the predicate it specifies. Consider the following examples:

- (9) A: What happens if I sell an eel?
 B: If you sell an eel, **you get 2.50 euros**.
 (10) A: When/Under which conditions do I get 2.50 euros?
 B: **If you sell an eel**, you get 2.50 euros.

Again, depending on the question different parts of the conditional are focused. In the first case, where the consequent is focused, a minimal model is one in which selling an eel results in nothing else than getting 2.50 euros. Hence, CP is not expected. With focus on the antecedent, as in dialogue (10), a minimal model is one where the only condition under which 2.50 euros are received is if an eel is sold. CP is expected to arise.

5 Experimental Investigations

The purpose of the experimental investigations was to test whether the predictions with respect to the influence of focus on CP are right. To investigate this six minimal context pairs were created that differed only in the question asked. The question was either of the type *what-if-p* or *when-q*. The answer to either question was a conditional of the form “If p, q”. Thus, questions of the type *what-if-p* put focus on the consequent of the conditional. Questions of the form *when-q* on the contrary, put the antecedent in a focus position. An alternative formulation for *when-q* would be *under-which-conditions-q* (cf. (10)). However, it was argued in the literature that questions with the Dutch equivalent are necessarily understood to ask for an exhaustive answer (cf. van Canegem-Ardijns & van Belle 2008: footnote 12). If this was also the case for the German

counterpart (the experiment was conducted in German) the questions would be useless to investigate the influence of focus. Thus we choose the *when-q* formulation. It is also known that for when-sentences a difference between a conditional and a temporal reading exists. In this study, we examined the conditional reading. To promote a conditional reading, we did not include terms referring to temporal aspects within the contexts of the test items. This was meant to ensure a conditional reading and with that bring about the effect that the answer “If p, q” rather than “When p, q” was not surprising or unintuitive. The task of the participants was then to judge in either case, whether the answer provided by the conditional was sufficient. Hence, felicity judgements rather than truth value judgements were collected. It is expected that a felicity judgement task (FJT) is more adequate to investigate whether implicatures were calculated (cf. Papafragou & Musolino 2003). The test items always specify two conditions that lead to a particular consequent (cf. 2). Thus participants are expected to rate the conditional answer as insufficient when CP occurred. If focus influences CP we expect that subjects rate the conditional answer more often as insufficient in the *when-q* condition than in the *what-if-p* condition.

Monika sells seafood on the market. She gets 1 euro for a crab, 2.50 euros for an eel, 15 euros for a lobster and 2.50 euros for a pike.

Kerstin, an employee of Monika, cannot remember the prices. Since she does not want to ask Monika again, she asks Sahra, who also works for Monika. Sahra knows the prices exactly.

<i>what-if-p</i>	<i>when-q</i>
Kerstin: What happens if I sell an eel?	Kerstin: When do I get 2.50 euros?

Sahra: If you sell an eel, you get 2,50 euros.

Did Sahra answer Kerstin’s question sufficiently? [Yes] [No]

Table 2: The table presents a sample item with both *what-if-p* and *when-q* questions.

For elicitation of the data a repeated-measures design was used with the question type as independent variable. Thus subjects were confronted with both types of questions, but never for the same context. The investigations were conducted in paper-and-pencil form and the questionnaire consisted of six test items and six fillers. Three out of the six test items contained *what-if-p* questions and the other three *when-q* questions. Each questionnaire contained three fillers that tested whether participants were in general able to understand conditionals as answers to questions. These *conditional fillers* were very similar

to the test items, but they specified only one antecedent for each consequent. Within a true conditional filler, the conditional answer corresponds to what has been described in the context. False conditional fillers provide the wrong antecedent for a consequence. Furthermore the questionnaires contained three fillers which tested whether subjects were sensitive for exhaustification. Like the test items, these fillers specified two conditions which will lead to the same consequent. Contrary to the test items, the answer in the dialogue was not given in conditional form, but was a simple statement starting with *only*. Answers of true *exhaustification* fillers provided both conditions that were specified by the context. The false exhaustification filler gave only one of these. Thus, if participants are in fact sensitive to exhaustification they should rate the response sufficient in the *true* condition and insufficient in the *false* condition. The partitioning was balanced over all six fillers, so that in total three were of the *true* condition and three of the *false* condition. Two versions of the questionnaire were used and items and fillers were presented in a pseudo-random order. No more than two test items were presented in a row. The same was ensured for the fillers. The second version of the questionnaire contained in each case the other condition of items and fillers and in the inverse order as in version one.

5.1 Procedure

The experiment was conducted in German. Participants received a paper copy of the questionnaire with written instructions on the first page. Before the questionnaires were handed out, participants also received oral instructions that summarized the written ones. Within the questionnaire, three items or fillers were printed on one page. At the end of the questionnaire participants were asked to provide some information about their background knowledge of logic and pragmatics. Additionally, space for comments was provided. The participants took about 15 minutes to fill out the questionnaire.

5.2 Participants

Participants were 50 students from an introductory linguistics class of the University of Frankfurt. Two subjects were excluded, prior to analysis, because they did not provide correct answers to three or more fillers. Hence, the data of 48 participants was evaluated.

5.3 Results

Negative answers, signaling CP, occurred in 89.2 % of the *when-q* contexts and in 16.3 % of the *what-if-p* contexts. An analysis of variance showed that over items as well as over subjects there was a main effect of question type (*when-q* vs. *what-if-p*) on the occurrence of CP, $F(1,5) = 145.93$, $p < 0.001$ and $F(1,47) = 309.93$, $p < 0.001$ respectively. Most participants also showed

a consistent behaviour over all test items. 61 % of the subjects labeled the answer as insufficient only with *when-q* and never with *what-if-p* questions. 33 % labeled the answers more often as insufficient in *when-q* contexts than in *what-if-p* contexts. Only 3 participants said that the answers are equally frequent insufficient in both question conditions. Moreover, nobody's answers signaled that CP occurred more often or even exclusively in *what-if-p* contexts.

5.4 Discussion

The results clearly indicate that focus has an effect on the occurrence of CP. This effect was highly significant over subjects as well as over items. We found that the percentage of *no*-answers, signaling CP, was high in *when-q* contexts and comparably low in *what-if-p* contexts. These results indicate that von Finchel's theory, together with Schulz & van Rooij's account on exhaustive interpretation, is adequate to account for CP. Focus seems to be the decisive factor for the occurrence of this inference.

However, the results are important in another respect. Zondervan (2009) investigated the effect of focus on the SI associated with *or* (the inference from "A or B" to "A or B but not both"). He also found more SIs when *or* was in a focus-position but the effect, though significant, was much smaller than expected. Zondervan found that the SI occurred in 77 % of the focused cases, versus 51 % of the non-focused cases. In contrast, the present study reflects a partitioning closer to the expected one. The main problem of Zondervan's results was the high amount of implicatures in the non-focus condition. This could be due to the fact that a different paradigm was used. As described, the present study used a FJT while Zondervan used a truth-value judgement task (TVJT) in his study. Within the TVJT, participants had to label the target sentences *true* or *false*. The underlying assumption is that a sentence like "A or B" is considered *false* in the case that the subjects calculated the SI ("A or B but not both") and indeed A and B holds. However, this appears to be rather a strong claim. The statement is clearly inappropriate if the SI is calculated but does that lead to falsehood? To judge this could be a problematic task for participants, so that the responses might not reflect natural understanding. Judging whether an answer is suitable or sufficient might be more natural. Additionally, labeling an utterance as *false* does not automatically mean that an implicature arose. As Zondervan (2006) discusses, a *false* answer can be interpreted in two ways: Either the subject calculated the implicature and thus labels the target sentence *false* in a situation where both A and B were the case. Or alternatively, the subject did not calculate an implicature but noticed that in the situation where both A and B are the case *and* would be more suitable than *or* and thus labels the target sentence *false*. So in addition to the problem that labeling sentences

true or *false* might be unnatural for participants, interpreting the results is difficult. Furthermore, the different results for the non-focus condition in this and Zondervan's study could be due to different properties of *or* and *if*. It could be the case that for *if* the occurrence of QIs is more dependent on focus, while *or* also elicits the inference regularly when not located in a focus position (for whatever reason). To investigate which of these two factors was decisive for the different results in Zondervan's and the present study a follow-up experiment was conducted.

6 Follow-up Experiment

The follow-up experiment replicated the first experiment with the exception that a TVJT was used. If this experiment yields the same results as the first experiment we can conclude that *if* and *or* behave differently when located in focus. If however, Zondervan's results are replicated we can conclude that the different paradigm was responsible for the response pattern.

6.1 Material

Test and control items were the same as in the previous experiment, with the only difference that the question "Did X answer Y's question sufficiently?" was changed to "Is the answer of X true?" This was done for test as well as control items.

6.2 Procedure

The procedure, the instructions, oral and written, and the design were identical to the first experiment.

6.3 Participants

36 students from an introductory linguistics class of the University of Frankfurt participated in the experiment. Two participants were excluded prior to analysis, because they stated having substantial knowledge of implicature theory. Additionally, one participant was excluded because she did not provide correct answers to three or more fillers. Thus, the data of 33 participants was evaluated.

6.4 Results

Within *when-q* context negative answers, signaling CP, occurred in 56.6 % of the cases and they occurred in 11.1 % of the *what-if-p* contexts. The effect was again significant over subjects, $F(1,32) = 58.175$, $p < 0.001$, as well as over items, $F(1,5) = 50.845$, $p < 0.001$. 52 % of the participants labeled the answer only in *when-q* contexts as insufficient. One half of the remaining participants labeled more answers insufficient in *when-q* contexts. The other half said that the answers are equally often insufficient in *when-q* and *what-if-p* contexts.

Nobody's answers signaled that CP occurred more often or even exclusively in *what-if-p* contexts.

6.5 Discussion

The results indicate once more that focus influences the emergence of CP and furthermore that this is independent of the paradigm used. The amount of CP was again significantly higher in the *when-q* condition than in the *what-if-p* condition. This holds over subjects as well as over items. The percentage of *no*-answers, indicating CP, is overall lower when a TVJT is used compared to the FJT. This may either indicate that felicity judgements are indeed more adequate to detect implicatures or that the FJT overestimates the amount of implicatures. However, what is interesting with respect to Zondervan is that using a TVJT also lowers the percentage of CP in the non-focus condition. This signals that the high percentage of implicatures which Zondervan found in the non-focus condition was due to properties of *or* rather than due to the TVJT.

7 Conclusion

This contribution provides evidence that von Stechow's (2001) theory together with the account of Schulz & van Rooij makes right predictions for the occurrence of CP. The experiments showed that focus influences whether an answer is interpreted exhaustively, i.e. whether CP arise. The amount of CP was significantly higher when the antecedent was in a focus position. We also found that the high amount of implicatures which Zondervan found in his study on *or* in non-focus conditions was probably not due to using a TVJT rather than a FJT. For the case of *if* the amount of CP was even lower in the non-focus condition when a TVJT was used.

As mentioned in the literature, not all implicatures behave the same (cf. Papafragou & Musolino (2003) on numerals and other scalar terms, Geurts & Pouscoulous (2009) on different embeddings and Chemla (2009) on scalar terms and free choice inferences). Geurts (2010: 122 ff.) makes an interesting observation with respect to potential differences of SIs and QIs that seems relevant for the different behavior of *or* and *if*. Even though Geurts proposes one account to derive both inferences he clearly distinguishes between the two types of implicatures. While we have a well-defined, closed set of potential alternatives for SIs such as, e.g., *all*, *most* and *many* for the scalar term *some*, this is not the case for QIs as those involved in exhaustive interpretation. For an answer like "Robin and Hilary left the party early" (example (6a)) no such clearly defined set of alternatives exists. We rather have to deal with an open-ended, possibly infinitely large set of alternatives. This core difference seems to play

an important role for the different experimental results found by Zondervan and my study. Zondervan investigated the SI associated with *or*, so a well-defined set of alternatives was present, namely *and*. Since CP as analyzed here is not a SI but a QI, an open set of alternatives exists which cannot be clearly defined. This difference appears to be crucial for the reasoning of the participants. As Zondervan had argued, labeling a target sentence *false* might not mean that the SI was calculated but only that the stronger lexical item would be more adequate. So maybe the amount of SIs is overestimated in the non-focus condition and possibly also in the focus condition. It is reasonable that with a finite and reasonably small set of alternatives, participants think about whether one of the alternative statements might have been more suitable. This could lead to the high amount of *false* ratings in the non-focus conditions. Within the present experiments on CP there are no concrete alternatives which could detract the participants and lead to a large amount of *no* answers in the non-focus condition. So answering *no* seems to be more clearly traceable to the emergence of the QI.

If this difference was indeed crucial for the different results for *or* and *if*, we could conclude that for the case of *or* the SI is also not elicited regularly in the non-focus condition. Rather the clear awareness of the better alternative statement *and* interferes and leads the subjects to conclude that the sentence is false. Furthermore, this predicts that items that are assumed to produce SIs (e.g., *some*) reproduce Zondervan's results, whereas items associated with QIs (e.g., the exhaustive interpretation in (6a)) should reproduce the results of the current study.

References

- Atlas, Jay David & Stephen C. Levinson. 1981. If-clefts, informativeness, and logical form. In Peter Cole (ed.), *Radical pragmatics*, 1–61. Academic Press.
- van der Auwera, Johan. 1997. Pragmatics in the last quarter century: The case of conditional perfection. *Journal of Pragmatics* 27. 261–274.
- van der Auwera, Johan. 1997b. Conditional perfection. In Angeliki Athanasiadou & Rene Dirven (eds.), *On conditionals again*, 169–190. John Benjamins Publishing Company.
- Boër, Steven E. & William G. Lycan. 1973. Invited inferences and other unwelcome guests. *Papers in Linguistics* 6. 483–505.
- van Canegem-Ardijns, Ingrid & William van Belle. 2008. Conditionals and types of conditional perfection. *Journal of Pragmatics* 40. 349–376.
- Chemla, Emmanuel. 2009. Universal implicatures and free choice effects: Ex-

- perimental data. *Semantics and Pragmatics* 2. 1–33.
- von Fintel, Kai. 2001. Conditional strengthening.
- Geis, Michael L. & Arnold M. Zwicky. 1971. On invited inferences. *Linguistic Inquiry* 2. 561–566.
- Geurts, Bart. 2010. *Quantity implicatures*. Cambridge University Press.
- Geurts, Bart & Nausicaa Pouscoulous. 2009. Embedded implicatures?!? *Semantics and Pragmatics* 2. 1–34.
- Grice, Paul. 1989. *Studies in the way of words*. Harvard University Press.
- Groenendijk, Jeroen A.G. & Martin J.B. Stokhof. 1984. *Studies on the semantics of questions and the pragmatics of answers*: University of Amsterdam dissertation.
- Horn, Laurence R. 1993. Economy and redundancy in a dualistic model of natural language. In Susanna Shore & Vilku Maria (eds.), *Sky 1993: Yearbook of the linguistic association of Finland*. 33–73.
- Horn, Laurence R. 2000. From if to iff: Conditional perfection as pragmatic strengthening. *Journal of Pragmatics* 32. 289–326.
- Horn, Laurence R. 2004. Implicature. In Laurence R. Horn & Gregory Ward (eds.), *The handbook of pragmatics*, Blackwell.
- Matsumoto, Yo. 1995. The conversational condition on horn scales. *Linguistics and Philosophy* 18. 21–60.
- McCawley, James D. 1993. *Everything that linguists have always wanted to know about logic ... but were ashamed to ask*. University of Chicago Press 2nd edn.
- Papafragou, Anna & Julien Musolino. 2003. Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition* 86. 253–282.
- van Rooij, Robert & Katrin Schulz. 2004. Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information* 13. 491–519.
- Rooth, Mats. 1996. Focus. <http://www.ims.uni-stuttgart.de/~mats/>.
- Schulz, Katrin & Robert van Rooij. 2006. Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy* 29. 205–250.
- Zondervan, Arjen. 2006. *The question under discussion focus condition for scalar implicatures*. Universiteit Utrecht MA thesis.
- Zondervan, Arjen. 2009. Experiments on QUD and focus as a contextual constraint on scalar implicature calculation. In Ulrich Sauerland & Kazuko Yatsushiro (eds.), *Semantics and pragmatics: From experiment to theory*, Palgrave Macmillan.