SCALAR IMPLICATURES *OR* FOCUS: AN EXPERIMENTAL APPROACH

Published by
LOT
Janskerkhof 13
3512 BL Utrecht
The Netherlands

phone: +31 30 253 6006
fax: +31 30 253 6406
e-mail: lot@uu.nl
http://www.lotschool.nl

Cover illustration:
Copyright © Pablo Picasso, *The Owl*, 1907, c/o Pictoright Amsterdam 2010.

ISBN: 978-94-6093-031-7
NUR 616

SCALAR IMPLICATURES *OR* FOCUS: AN EXPERIMENTAL APPROACH

Scalaire Implicaturen of Focus: een Experimentele Benadering
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. J.C. Stoof,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
vrijdag 28 mei 2010 des middags te 2.30 uur

door

Albert Jan Zondervan

geboren op 28 augustus 1980 te Amsterdam

Promotoren: Prof. dr. F.N.K. Wijnen
Prof. dr. H.E. de Swart

CONTENTS

ACKNOWLEDGMENTS

Whenever I received one of the dissertations of my colleagues at UiL-OTS, which are usually generously distributed, I always found it odd that the acknowledgements preceded the main text. I felt it was like thanking people before the actual accomplishment had been made, like an alpine climber at the foot of Mount Everest saying before taking off: "I would like to thank the people that helped me reach the top." Therefore, I decided my acknowledgements section would follow the main text. Obviously, I changed my mind. I guess the main reason for this is that after 10 months or so of writing in isolation in a basement, I realize how important the people were that helped me during my time as a PhD student, and that they deserve a prominent place in the dissertation.

Of course, I am writing this section in a hurry. I am one of those students that never finishes something until the very last moment (or preferably even shortly after that). Therefore, I will forget to mention many people that I am grateful to and I hope they know this despite the fact that they are not mentioned. A note to the Dutch people that I thank in this section: I hope you are not bothered by the fact that I thank you in English. I have to say I am glad I can write this in English, as to me, as a non-native speaker of English, it sounds a lot less sentimental than in Dutch.

First of all I would like to thank my supervisors, Frank Wijnen and Henriëtte de Swart. What I appreciated most about their supervision was their trust in my ability to design and carry out the project myself. They gave me complete freedom as to which topics to investigate and which experiments to run. I think it is a great virtue for a supervisor to not try to steer a student in a certain direction. Frank en Henriëtte always made me feel like I was the expert on the topic and I was calling the shots. However, they were always there when I needed them to help me proceed. I would also like to thank them for going along with my crazy last minute planning in the last stages, and apologize for making them read drafts of chapters over the weekend. I would also like to thank Tanya Reinhart, who was my supervisor in the first months of my project. We only had a couple of meetings but I learned a lot from them.

I thank Maaike Schoorlemmer and Martin Everaert for arranging everything that made it possible for me to have a successful career as a PhD student at UiL-OTS. I especially appreciate that I got to go to many interesting conferences and schools around the world, which helped me a lot in finding my way in the international world that is linguistics.

Besides my supervisors there were many people who took time out of their busy schedules to meet with me and talk about my research, for which I

am very grateful. These people are (in random order): Bart Geurts, Robert van Rooij, Petra Hendriks, Lyn Frazier, Chris Potts, Emmanuel Chemla, Gennaro Chierchia, Chungmin Lee, Daniele Panizza, Vincenzo Moscati, Edgar Onea, Henk Pander Maat, and many more I forget. A couple of people are missing from this list because they deserve a special note. The first is Øystein Nilsen. He is one of the two people who sparked my interest in scalar implicatures (together with Kriszta Szendroi). When I was a student in the masters program in Utrecht, Øystein taught us about this topic. I remember he spent at least an hour trying to explain that the sentence *The girl is beautiful* triggers the scalar implicature that she is not *stupendous*. However, I (and probably many of my Dutch classmates) did not understand anything of what he was saying, as I had no idea what the word *stupendous* meant, and was obviously too ashamed to ask. Eventually he must have gotten the point across, as I chose to further investigate this topic. Øystein and I had many meetings throughout my masters and PhD studies about my research and the possibility of linking it up with his ideas. Although it never really materialized in a paper I appreciated these meetings a lot. I admire Øystein for his enthusiasm and for spontaneously approaching students and meeting up with them to talk about their research. I think this type of enthusiasm contributes a lot to a research community such as UiL-OTS, even though it might be hard to measure.

The second person, or rather persons, that discussed my work with me and deserve a special note are Andrea Gualmini and Luisa Meroni. I would like to thank them for giving me a taste of what scientific collaboration is like. I'm sorry that the Question Under Discussion did not make it to the dissertation, but this is for explanatory reasons only: it turned out the term itself raises a lot of misunderstanding, which I wanted to avoid. But this does not mean I don't believe in it anymore.

The third person who discussed my work with me but was not in the list above is Napoleon Katsos. The reason is that he deserves both a thank you and an apology. I thank him for talking to me about my work and for his dissertation, which was extremely helpful in writing my own. I apologize for raising so many points of criticism about his experiments. However, I am sure he will be able to raise many more points of criticism about the experiments presented in this dissertation, which I of course beg him not to do.

Although I am not sure I am allowed to do this by the strict regulations of the Dutch system of PhD thesis assessment, I would like to thank the committee for reading and assessing the manuscript, which must have been quite a time-consuming matter: Bart Geurts, Michiel van Lambalgen, Ira Noveck, Robert van Rooij and Ted Sanders.

I spent much of my time as a PhD student designing and conducting experiments. I am very thankful to all the people who participated in these experiments, and all of my friends, family and acquaintances who managed to get their friends, family and acquaintances to fill out the web-based questionnaires with silly stories about people searching for marine animals while using the word *or* all the time. It was heartwarming to, after having sent out an email asking for participants, see my mailbox fill up with notifications of the questionnaire software that another participant filled out the experiment. I am also thankful to all the professors and fellow PhD students that let me interrupt their teaching to bother their students with a questionnaire: Frank Wijnen, Joost Zwarts, Dagmar Schadler, Sander van der Harst, Loes Koring, Huib Kranendonk, Hannah De Mulder, Roberta Tedeschi, Hans Van de Velde and many more I forget. I also thank Iris Mulders for creating, maintaining and helping me use the brilliant system for recruiting participants for the lab experiments. I furthermore thank Theo Veenker for creating and helping me use the questionnaire software WWSTIM, and I thank him, Iris and Warmolt Tonckens for providing technical lab support. Finally, I thank Hugo Quené and Loes Koring for recording the spoken stimuli for one of the experiments.

I was fortunate to have a big group of fellow PhD students at UiL-OTS and I would like to thank those that crossed my path over the years for being great colleagues: Frans Adriaans, Ana Aguilar, Anna Asbury, Natalie Boll-Avetisyan, Ivana Brasileiro, Bert Le Bruyn, Desiree Capel, Anna Chernilovskaya, Xiaoli Dong, Jakub Dotlacil, Sanne Elling, Lizet van Ewijk, Gaetano Fiorin, Berit Gehrke, Gianluca Giorgolo, Nicole Grégoire, Nadya Goldberg, Nino Grillo, Bettina Gruber, Sander van der Harst, Naomi Kamoen, Annemarie Kerkhoff, Cem Keskin, Arnout Koornneef, Loes Koring, Huib Kranendonk, Joke de Lange, Tom Lentz, Hannah De Mulder, Emilienne Ngangoum, Andreas Pankau, Min Que, Rianne Schippers, Marieke Schouwstra, Giorgos Spathas, Roberta Tedeschi, Christina Unger, Rosie van Veen, Nina Versteeg, Daphne van Weijen, Marie-Elise van der Ziel, Rob Zwitserlood, and everyone I forget. At the beginning of my project the institute had just moved to the building at Janskerkhof, and we made an effort to re-instate old PhD traditions like a common lunch and the Friday afternoon drinks (VrijMiBo). I did not participate much in the common lunch as I was too lazy to bring my own food, but the VrijMiBo was (and is, so I have been told) a big success. Sharing thoughts and frustrations over a beer at Café Hofman with people who are in the same process always gave me extra motivation for my work.

I thank a number of fellow students from the above list in particular. First of all, my roommates of Janskerkhof 13 room 0.05, by far the coolest room in the building (in both senses of the word *cool*). You guys in the attic

may have had airconditioning, but we had Arnout, Hannah, Roberta, Natalie, Loes and Nadya, an orange soccer ball and a candy grabber! I would also like to thank Nicole and Sanne for organizing the Uiltjesdagen 2007 with me, long live the Smoelenboek. I thank Hannah and Tom, together with Martin Everaert and Øystein Nilsen, for turning the UiL-OTS yearbook into 'Theoretical Validity and Psychological Reality', into 'Capturing Cognition' and eventually into 'The Linguistics Enterprise' with me. I thank the Dutch posse (Hannah, Tom, Frans, Loes and Arnout) for numerous lunches, drinks and dinners at places that apparently serve good vegetarian food. I hope many of these occasions will follow. Finally, I thank Hannah and Arnout for being my paranymphs.

At UiL-OTS I would also like to thank Rick Nouwen, Øystein Nilsen and Jakub Dotlacil for organizing and inviting me to the reading groups on semantics. I know I probably did not always seem that interested, but this was just a pose to cover up the fact that I often did not understand anything of the smart stuff they were talking about. I remember a session where they were going on about 'ACD', which I found out later was not a band name with a letter missing. I also thank Jacqueline van Kampen, for hiring me as a student-assistant when I was an undergraduate student, and for being the 'manager' of Achter de Dom and looking after me during the last months in the basement. Finally, I thank Frank Wijnen and Alexis Dimitriadis for giving me the opportunity to teach the tutorials of Intro to Psycholinguistics and Intro to Semantics.

During my time as a PhD student, there were also some quiet days. The liberal work ethos at UiL-OTS and the fact that I was not working in a project with other people sometimes resulted in the situation that I was alone in the office and the first words that came out of my mouth that day were directed at the cashier when I was getting groceries for dinner (usually: "Pinnen graag."). Obviously, this happened very regularly when I chose to be in solitary confinement in the basement of Achter de Dom to write up. What kept me going those days were the social events at night and in the weekends (planned and spontaneous) with friends, family, housemates and teammates, and I thank them all for that. I thank Connie and Hans for the numerous 'duwpotten' coffee at Brandmeesters, which made me get up and start work at an acceptable hour.

A very important thank you goes to my parents, Hans en Janne. They always let me make my own choices and supported me unconditionally in everything I did. Perhaps they are the main reason I became a linguist. When I went to college/university, they told me go and study something that I found interesting, not something that would later help me make a lot of money. Well, I did. I would also like to thank my sister Karin for supporting me. Although we studied different topics at different universities, we

discovered that our academic interests are very similar, but it is obvious she is the real academic talent of the family.

Finally, I thank Monique. I cannot describe how much more than the owl on the cover of this book you have contributed to this work and to my life.

CHAPTER 1: INTRODUCTION

1. Introduction: Meaning, truth conditions and the pragmatics wastebasket[1]

This dissertation is about meaning of language. Most people think of meaning as the relation between a word and a certain concept in the mind. However, meaning in natural language is not just that. We can use language to communicate about the world around us. Therefore, meaning is not just about concepts in our mind, but it is linked to the outside world. In that sense, meaning is *referential*. Consider for instance the meaning that is communicated when I utter sentence (1):

(1) My dog bit my cat yesterday.

In sentence (1) I am not just activating the concepts of *dog*, *cat*, *biting* and *yesterday*, I am referring to actual things in the outside world: my dog, my cat, an actual biting incident and an actual day. Most scholars in the field of semantics, which is the subdiscipline of linguistics which is concerned with meaning, work within this referential theory of meaning. The reference of natural language expressions is not limited to concrete objects, but includes more abstract things such as relations, properties and situations. Even the meaning of words that refer to non-existent things such as *unicorn*, or sentences about non-actual situations (*I wish I was the King*) can be captured in a referential theory of meaning once we abstract over possible worlds.

A referential theory of meaning has to be more than just the reference of words to things in the real world (or a hypothetical world), otherwise (2) would have the same meaning as (1):

(2) My cat bit my dog yesterday.

Apparently, it makes a difference for the meaning of a sentence how the elements are combined, i.e. the structure of the sentence has to be taken into account. This is one of the reasons why semantics is mainly concerned with the meaning of *sentences*, instead of the meaning of words. Another reason for this is that the sentence is taken to be the smallest unit which communicates a 'complete thought'. Nevertheless, most scholars adhere to

[1] This section in based on chapters 1 and 2 of the semantics textbook of De Swart (1998) (but any errors are all the current author's). The interested reader is referred to this source for a more elaborate introduction to truth-conditional semantics.

the *Principle of Compositionality of Meaning* (usually attributed to the German philosopher Gottlob Frege), which says the meaning of the whole is fully determined by the meaning of the parts and the way they combine. As the parts of a sentence are words, the meaning of a sentence has to be somehow built up from the meaning of the words of the sentence and the rules by which these are combined. This introduces an important restriction on the type of meaning that is described in semantics: it is the 'literal' meaning which follows from the meaning of the words and their relation to each other. Crucially, this meaning does not include the parts of meaning that arise through the specific context or situation in which the sentence is uttered. In order to avoid confusion over the different uses of the word *meaning*, we use the term *sentence meaning* for this literal meaning of the sentence, and *utterance meaning* or *speaker meaning* for the 'complete' meaning of an utterance, including the meaning derived from contextual and situational information.

If we take the view of sentence meaning as referring to things in the world, we can capture the meaning of a sentence by defining what the world should look like in order for it to be true. Irrespective of whether someone actually uttered (1), we know what it means because we know exactly when it is false and when it is true. 'True' and 'false' are called *truth values*, and the conditions under which a sentence is true are called the *truth conditions* of the sentence. Crucially, we are able to 'calculate' the truth conditions of a sentence from the meaning of the words and the way they are put together. This follows from the fact that for every new sentence we hear, irrespective of whether we ever heard this sentence before or whether we know if it is true or false (its truth value), we know what its truth conditions are.

This truth-conditional view of sentence meaning allows us to define the meaning of certain words or phrases which combine with whole sentences, in terms of their interaction with the truth conditions of those sentences. Two examples of such words are *and* and *not*. Consider for instance (3) and (4):

(3) My dog did not bite my cat yesterday.

(4) My dog bit my cat yesterday and my cat bit my dog yesterday.

In (3), we combine *not* with the sentence (1). As a result, the truth conditions of (3) are the opposite of (1): whenever (1) is true, (3) is false and vice versa. Therefore, we can define the meaning of *not* in terms of this reversal of truth conditions. Similarly, the truth conditions of (4) are based on the truth conditions of the two sentences connected by *and* (in this case (1) and (2)). *And* determines that if and only if both of them are true, the whole sentence (4) is true. In the next section we will see that a similar truth-conditional case

can be made for the meaning of *or*. The interaction between this truth-conditional meaning of *or* and the interpretation of *or* in natural language is the topic of this dissertation.

Due to the truth-conditional approach to meaning and the compositionality principle that sentence meaning has to be derivable from the meaning of its parts, semantics idealizes over the specific situation in which a sentence is uttered and the specific context it appears in. In order to capture the truth conditions of (1), we do not want to include the specific speaker who utters (1), as that would result in a different definition of the sentence meaning of (1) for every speaker. This seems to miss the point that (1) has a meaning irrespective of who utters it. However, it is obvious that in normal conversation, it is meaningful who uttered (1), as this tells us something about how the world is. Parts of utterance meaning which follow from the specific situation in which a sentence is uttered are the topic of study of the subdiscipline of linguistics called *pragmatics*. Next to utterance meaning originating from who the speaker and the addressee is, pragmatics covers all kinds of non-literal meaning derived from the specific situation, such as meaning of ironic or sarcastic use of language, or the meaning that arises when language is used to perform a certain act. The most famous illustration of the latter phenomenon is the difference between the literal meaning of *Can you pass the salt?* and the request that is made by it.

There are also parts of utterance meaning stemming from the linguistic context in which a sentence is uttered that cannot be captured by truth-conditional semantics of sentences. An example of this is the reference of pronouns over sentences. Consider (5):

(5) My dog bit my cat yesterday. It was hurt badly.

There is nothing in the truth conditions of *It was hurt badly* which tells us whether *it* refers to *my dog* or to *my cat* in the previous sentence. There is also no way to derive this from the meaning of the words of the second sentence in a compositional way. Therefore, this aspect of utterance meaning also falls outside the scope of truth-conditional semantics and is considered part of pragmatics. There is also a type of utterance meaning of material in the sentence itself that cannot be captured by truth-conditional semantics, see (6) and (7):

(6) I regret having bought a dog.

(7) I do not regret having bought a dog.

From (6), it becomes clear that I bought a dog. This is however not part of the truth conditions of (6), as adding *not* to it (as in (7)) does not make it false. From (7) we can still conclude that I bought a dog. The meaning that I bought a dog is considered to be a background assumption of (6), not something that is asserted by it. This part of utterance meaning is called *presupposition* and it is also considered to belong to pragmatics.

The topics studied in pragmatics seem to be selected by the criterion that truth-conditional semantics cannot account for them. Therefore, a metaphor that is regularly used is that pragmatics is the *wastebasket* of semantics, or of linguistics in general. As a result of the idealization of meaning to truth-conditional semantics, researchers have been able to develop very exact theories of sentence meaning. The downside of this approach is that the parts of utterance meaning that ended up in the pragmatics wastebasket have been studied far less systematically. This dissertation is concerned with one of the types of utterance meaning from the pragmatics wastebasket, conversational implicatures. Specifically, it will consider the scalar implicature (a subtype of conversational implicature) associated with *or*. I will introduce conversational implicatures and scalar implicatures in the next section, and show how they are also the result of the desire to keep truth-conditional semantics clean. In section 3 I will introduce the specific topic and the goals of the thesis. The outline of the chapters and a preview of the most important results is given in section 4. Finally, I summarize the research program in section 5.

2. The starting point of the thesis: Scalar Implicatures and *or*

2.1 Conversational Implicatures

One type of utterance meaning that is not part of the truth-conditional meaning of sentences and cannot be retraced to the meaning of its parts, is meaning that arises because we consider the speaker who uttered the sentence to be cooperative. This assumption is known as the *Cooperative Principle*, put forward by Grice (1967):[2]

> The Cooperative Principle:
>
> Make your contribution such as required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

[2] Grice's William James Lectures of 1967 called *Logic and Conversation* were first published in 1975, but for clarity I will refer to them as Grice (1967).

Grice made the idea of the Cooperative Principle explicit by listing four *maxims of conversation*, guidelines that according to him every speaker is expected to follow in normal conversation:

The maxim of Quality:

Try to make your contribution one that is true, specifically:

(i) Do not say what you believe to be false.
(ii) Do not say that for which you lack adequate evidence.

The maxim of Quantity:

(i) Make your contribution as informative as is required for the current purposes of the exchange.
(ii) Do not make your contribution more informative than is required.

The maxim of Relation/Relevance:

Make your contribution relevant.

The maxim of Manner:

(i) Avoid obscurity.
(ii) Avoid ambiguity.
(iii) Be brief.
(iv) Be orderly.

Apparent violations of these maxims give rise to a part of utterance meaning that Grice called *conversational implicature*. Consider for instance example (8) from Levinson (1983):

(8) A: What on earth has happened to the roast beef?
B: The dog is looking very happy.

At first glance, B's answer seems to violate the Cooperative Principle and specifically the Maxim of Relation/Relevance, as the statement about the dog seems to be irrelevant to the question. However, the addressee will typically try to interpret B's statement in such a way that B is cooperative. Therefore, A will assume that B's statement is somehow relevant, and using the context she might infer the meaning paraphrased in (9):

(9) I don't know what happened to the roast beef, but the dog might have eaten it because it looks very happy.

However, there is nothing in the meaning of the words of *The dog is looking very happy* or the way in which they are combined that could account for the first part of (9) (the part before *because*). If this were so, then every occurrence of *The dog is looking very happy* would give rise to this meaning. This is obviously not the case, as (10) shows:

(10) A: How is the dog today?
B: The dog is looking very happy.

Thus, compositional semantics cannot derive the inference that the addressee made in (9). According to Grice and his followers it is a conversational implicature, derived from the assumption that the speaker is being cooperative and adhering to the maxims of conversation.

Conversational implicatures do not always have to be as situation-specific as (8). Consider (11), also from Levinson (1983):

(11) The flag is white.

The conversational implicature here is that the flag is *entirely* white. This implicature can be derived by Grice's Maxim of Quantity (i) (*Make your contribution as informative as is required*). If it were the case that the flag had other colors too, say we are talking about the Russian flag which is white, blue and red, a cooperative speaker would have mentioned the other colors too, because that would have been more informative. Therefore, the addressee will take (11) to mean that the flag is entirely white. If the information that the flag is entirely white was part of the compositional semantic meaning of the sentence, this meaning would also be part of the meaning of sentence (12), which would then be a contradiction:

(12) The flag is white, blue and red.

This dissertation is concerned with conversational implicatures of this latter type: the type where an addressee lands on a certain interpretation of a sentence based on the fact that the speaker did not use another, more informative expression. The subtype of conversational implicatures that will be studied is called *scalar implicatures*. I will discuss these in the next subsection.

2.2 Scalar Implicatures

Some linguistic expressions form a scale with other expressions. The most intuitive examples of these are adjectives like *terrible, bad, okay* and *excellent*. These four words reflect points or ranges on a scale of 'goodness'. Similarly, *small, medium-(sized), large* and *huge* are on a scale of 'bigness'. Using a certain item of a scale can give rise to a specific type of conversational implicature. Consider for instance (13):

(13) Q: How was the concert last night?
A: It was okay.

The answer in (13) implies that the concert was not excellent. This seems to be the result of the fact that *okay* and *excellent* are on the same scale, and *excellent* is higher on this scale than *okay*. Therefore, as the speaker picked the weaker element on the scale, the hearer concludes that the sentence with the stronger item (*excellent*) does not hold. This inference is a specific type of conversational implicature called a *Scalar Implicature* (SI). Before I get to the exact derivation of this inference, I will explain why this part of the utterance meaning is considered to be an implicature, and not part of the truth-conditional meaning of the answer in (13A).

A first indication that the inference that the concert was not excellent is not part of the literal sentence meaning of (13A), is that this conclusion is cancelable. We can cancel the inference by adding material to the contrary, without creating a contradiction. For instance, the answerer in (13) could have said (14):

(14) It was okay. In fact, it was excellent.

The continuation in (14) does not seem to give rise to a contradiction. This is one of the hallmarks of conversational implicatures, the property of *cancelability*. Notice that we cannot do the same thing with the lower bound of *It was okay*, see (15):

(15) It was okay. # In fact, it was bad.

The '#' symbol indicates that the second sentence is not a good continuation of the first. The continuation seems to create a contradiction, which indicates it goes against the truth-conditional meaning of the first sentence. The contrast in (14) and (15) shows that the lower bound of *okay* (that the concert was *at least* okay) is part of the truth-conditional meaning of *okay*, but the upper bound (that the concert was not excellent) is not.

The second argument that the inference that the concert was not excellent is not part of the truth-conditional meaning of *It was okay* is a bit more complex. Consider what happens when we put (13A) under negation (i.e. we add *not*), see (16):

(16) The concert was not okay.

As illustrated in (1) and (3) above, *not* reverses the truth conditions of the sentence with which it combines. Compare for the sake of the argument the meaning of *okay* and *excellent* to grades on a 10-point scale. We could define *excellent* as a grade between 8 and 10. Let us also assume that in order for something to be *okay* it has to at least have the grade 6. If *okay* in (13) means *okay but not excellent*, it means the concert was between a 6 and an 8. If this is the meaning of *okay*, then (16) should mean that the concert was not between a 6 and an 8. But that is not what (16) means. It is clear that (16) means the grade of the concert was not between 6 and 8, but also not between 8 and 10. In other words, *okay* in (16) covers the range from 6 to 10. It seems *okay* has two different meanings: in (13) it means '6-8' and in (16) it means '6-10'. We can account for this apparent ambiguity by assuming the *not excellent* part in (13), which rules out the range 8-10, is not part of the meaning of *okay*, but comes about by a scalar implicature. On this account, we can keep one basic meaning of *okay*, the meaning that ranges from 6 to 10, and still explain why intuitively it refers to the range from 6 to 8 in (13).

People often have a hard time accepting this explanation, because in their perception, the meaning of *okay* as the range from 6 to 8 is the basic meaning. This is because SIs like these occur very frequently and it is tempting to take the meaning of the simple sentence (without for instance *not*) as the basic case. However, there are many environments which obviously trigger the broader meaning. Consider for instance (17), (18) and (19):

(17) If the concert was okay, you should write a review of it.

(18) Every concert that I went to which was okay, was in Wembley Stadium.

(19) None of the concerts that I went to was okay.

Sentence (17) tells you that you should also write a review if the concert was excellent, (18) entails that if there were any concerts that the speaker went to that were excellent, these were also in Wembley Stadium, and (19) says that

the speaker has never been to an excellent concert either. So in all these cases, *okay* seems to have the broad '6-10' meaning. These examples support the claim that this meaning of *okay* is basic, and the '6-8' meaning in (13) is the result of an SI.

However, now the question arises why the SI does arise in (13), but not in (16)-(19). To explain this, we have to consider the notion of *informativity*. On the view that the basic meaning of *okay* is the 6-10 meaning, and *excellent* means something like 8-10, it seems that *excellent* is more informative than *okay*, as it puts a stronger restriction on how the real world is. After all, *excellent* also rules out that the concert was in the 6-8 range, which is allowed by *okay*. As Horn (1972) proposed, this notion of informativity can be described in terms of *entailment*: If sentence A entails sentence B it has to be the case that whenever A is true, B is also true. This is the case for the sentences *It was excellent* and *It was okay*: whenever a concert was excellent, it was also okay. However, the reverse does not hold. We can therefore define informativity in terms of *asymmetrical* entailment: if one expression asymmetrically entails the other, the first is more informative than the second. Entailment scales which can give rise to SIs and come about by this definition are therefore called *Horn-scales*. Here are some examples from Horn (1972), as listed in Levinson (1983) (p. 134).[3] These scales are usually represented between '< >'. It is my choice to put the strongest item on the right side and the weakest on the left.

(20) <few, some, many, most, all>
<good, excellent>
<warm, hot>
<sometimes, often, always>
<want to *V*, try to *V*, succeed in *V*-ing>
<possibly *p*, *p*, necessarily *p*>
<possible that *p*, probable that *p*, certain that *p*>
<may, should, must>
<cool, cold>
<like, love>

The list in (20) and the knowledge that there are many more of these scales shows how widespread the phenomenon of scalar implicature is.

[3] I left out three examples which are mentioned in Levinson (1983): <or, and> which will be discussed below, the scale of numerals because their status as SI-triggers is debated and the <not all, none> scale as it might interfere with my discussion of DE-environments.

Now that we have defined informativity in terms of asymmetrical entailment, we can explain how exactly the SI arises in (13) and why it is absent in (16)-(19). Example (13) is repeated here as (21) for convenience:

(21) Q: How was the concert last night?
A: It was okay.

Based on the Cooperative Principle and specifically the Maxim of Quantity (i) (*Make your contribution as informative as is required*), the addressee in (21) will assume that the speaker will use the most informative item on the scale which is compatible with her beliefs. Therefore, as the addressee knows that *It was excellent* would have been more informative, she can conclude that the speaker did not consider the concert to be excellent. This kind of reasoning can be represented step-by-step as in (22). A derivation such as (22) below is often called a *nonce derivation*.

(22) i. The speaker used the scalar term *okay*.
ii. The speaker could have uttered the same sentence with the scalar term *excellent* instead of *okay*, which would have been stronger/more informative (because the sentence with *excellent* entails the sentence with *okay*).
iii. The sentence with the stronger scalar term *excellent* would have also been relevant.
iv. The speaker is trying to be as informative as possible (she is obeying the Maxim of Quantity).
v. Apparently, the speaker does not have evidence for the sentence with *excellent*.
vi. The speaker is well informed.
vii. Therefore, it is likely that the speaker considers the sentence with *excellent* to be untrue.

If the addressee in (21) applies this reasoning, she will conclude that *The concert was excellent* is false and therefore she will take it that the concert was okay but not excellent. Whether language users actually go through these steps and to what extent this is a conscious process is debated in the literature, but many scholars agree that the reasoning in (22) provides the key ingredients for calculating SIs.

The reason the SI is not triggered in (16)-(19) is that these are all *Downward Entailing* (DE) environments, in which the informativity relation between *okay* and *excellent* is reversed. Consider for instance (16) (*The concert was not okay*). This rules out that the grade of the concert was anywhere in the range 6-10. Hence, it must have been somewhere in the

range between 1 and 6. The same sentence with *excellent* (*The concert was not excellent*) only rules out the range 8-10, leaving the range between 1 and 8 as possible. Therefore, the sentence with *okay* is in this case more restrictive and thus more informative. In terms of entailment: *It was not okay* asymmetrically entails *It was not excellent*. This is why the SI does not arise in (16): using *excellent* instead of *okay* would not have led to a more informative statement. Therefore, step (ii) of (22) will not go through in a DE-environment. For (17)-(19) a similar case can be made. I now turn to the specific SI that this dissertation is concerned with: the SI of *or*.

2.3 The scalar implicature associated with *or*

In section 1 I discussed truth-conditional semantics and how we can capture the meaning of connectives like *not* and *and* in terms of their truth-conditional contribution. Traditionally, the meaning of *or* is also described in terms of truth conditions.[4] In truth-conditional semantics, *or* is taken to denote logical disjunction ($\vee$), which yields a true sentence if at least one of the sentences connected by it is true. This corresponds to what is called *inclusive-or*, as it also yields a true sentence if both sentences it connects are true. This meaning can be observed in sentences like (23):

(23) I'd be surprised if John took a shower or brushed his teeth.

Clearly, if I utter (23) I would also be surprised to hear that John did both things. Therefore, *or* in (23) means *and/or*, i.e. at least one of the two sentences connected by *or* has to be true, and possibly both are. Hence, *A or*

[4] *Or* is also often used in natural language in ways which do not correspond to the truth-conditional meaning. Consider for instance the following passages from the movie 'The Big Lebowski' (written by Joel and Ethan Coen).

(i) "I'm the Dude. So that's what you call me. That, or Duder. His Dudeness. Or El Duderino, if, you know, you're not into the whole brevity thing."

(ii) "Oh, Cynthia's Pomeranian. Can't leave him home alone or he eats the furniture."

In (i), *or* has a meaning similar to the adverb *alternatively*. In (ii), *or* does connect two sentences, but it adds another meaning than that one of them is true. The meaning of *or* in (ii) is similar to the adverb *otherwise*. As these quotes illustrate, *or* is also often used as an adverb instead of a truth-conditional connective. However, these uses of *or* are not the type that I will be concerned with in this dissertation as they do not give rise to SIs.

B is interpreted as *A or B or both* on this reading. However, sometimes *or* seems to have a different meaning. Consider for instance (24):

(24) John took a shower or brushed his teeth.

The most natural interpretation of (24) is that John *either* took a shower *or* brushed his teeth, i.e. he did one of the two things. On the basis of such sentences, we could propose that the truth-conditional meaning of *or* is that *exactly one* of the two sentences it connects must be true. On this reading, *A or B* actually means *A or B but not both*. This reading is often called the *exclusive-or* reading, as it rules out the possibility that both sentences are true. It seems that *or* gives rise to an ambiguity between the inclusive-*or* reading and the exclusive-*or* reading.

However, as first proposed by Horn (1972), the two readings of *or* can be explained in the same way as the apparent ambiguity of *okay* above: by an SI. The basic truth-conditional meaning of *or* is taken to be the one illustrated in (23), inclusive-*or*, corresponding to the disjunction operator in formal languages, usually represented with the symbol $\vee$. On this inclusive meaning, *or* is on an entailment scale with *and*. To see this, consider the four situations in (25) below, which cover the range of possibilities with the two sentences which are connected by *or* in (24).

(25) Situation 1: John both took a shower and he brushed his teeth.
Situation 2: John took a shower but he did not brush his teeth.
Situation 3: John brushed his teeth but he did not take a shower.
Situation 4: John did neither.

If we interpret sentence (24) on the inclusive reading of *or*, Situations 1-3 are possible. However, the same sentence with *or* replaced by *and* (*John took a shower and brushed his teeth*) yields only Situation 1 as a possibility. So every situation in which *A and B* is true (where *A* and *B* are placeholders for sentences, or in semantic terms *propositions*), *A or B* is also true. Hence, *A and B* asymmetrically entails *A or B*. Therefore, using *A or B* can trigger the SI that *A and B* is not the case. We can provide a similar nonce derivation for (24) as we did for (13) above. See (26):

(26) i. The speaker used the scalar term *or*.
ii. The speaker could have uttered the same sentence with the scalar term *and* instead of *or*, which would have been stronger/more informative (because the sentence with *and* entails the sentence with *or*).
iii. The sentence with the stronger scalar term *and* would have also been relevant.
iv. The speaker is trying to be as informative as possible (she is obeying the Maxim of Quantity).
v. Apparently, the speaker does not have evidence for the sentence with *and*.
vi. The speaker is well informed.
vii. Therefore, it is likely that the speaker considers the sentence with *and* to be untrue.

If the addressee applies the reasoning in (26) to (24), she will conclude that *John took a shower and brushed his teeth* is false, yielding the exclusive-*or* reading *John took a shower or brushed his teeth but not both*. Similar to (16)-(19) above, the SI does not arise if *or* is embedded in a DE-environment, as its entailment relation with *and* is reversed. This is exemplified in (23), where we keep the basic inclusive-*or* reading.

The cancellation pattern also supports the SI-view of the meaning of *or*. Consider (27) and (28), which show a similar pattern as (14) and (15) for *It was okay.*

(27) John took a shower or brushed his teeth. In fact, he did both.

(28) John took a shower or brushed his teeth. # In fact, he did neither.

This contrast indicates that the part of the meaning of (24) that John did not do both is not part of the truth-conditional meaning, as it can be canceled, while the meaning that at least one of the two has to be the case is truth-conditional, as negating it leads to a contradiction.

The main advantage of the SI-view of *or* is that we can stick to the view that *or* has one basic meaning (inclusive-*or*) and that exclusive-*or* is derived from this meaning through pragmatic means (the SI). This is another case where the pragmatics wastebasket keeps truth-conditional semantics clean.

3. Topic and goals of the thesis

As I explained above, pragmatics has traditionally been considered as the wastebasket of semantics and of linguistics in general. Thanks to the pragmatics wastebasket, scholars working on truth-conditional semantics have been able to make great progress in modeling meaning in natural language. However, as a result of the focus on truth-conditional sentence meaning, the aspects of utterance meaning that have been moved to the pragmatics wastebasket have often been left out of consideration, and have not been accounted for in such a well-structured way yet. Parts of utterance meaning such as conversational implicature have not been described anywhere near as systematic as parts of utterance meaning which belong to compositional semantics. Many pragmatic theories are successful in explaining after the fact how a particular example gave rise to a certain meaning, by relying on general principles which are formulated in an inexact way, while theories that make clear predictions are quite rare. This is a sharp contrast with theories in compositional semantics, which are usually very explicit. This dissertation is an attempt to add to a more structured account of pragmatic phenomena, by distilling testable predictions from pragmatic theory and testing these predictions experimentally.

It is therefore no trivial choice to focus on scalar implicatures. Scalar implicature is considered to be one of the most robust pragmatic phenomena in the sense that it is less dependent on specific circumstances than many other pragmatic inferences. Therefore, it is a promising topic in the search for regularities within pragmatics. However, at the current point of research in pragmatics, even for these inferences it is far from clear when they arise and when they are absent. As I will illustrate in the next section, it is well known that whether SIs arise depends on the context in which a sentence with a scalar term like *or* occurs. However, very little is known about what the properties of the context are that are responsible for this. The question when one of the most robust types of pragmatic inference - scalar implicature - arises, should be a fundamental question for a more explicit theory of pragmatics. Therefore, I will investigate this question in this thesis.

Just as it is no trivial choice to focus on scalar implicatures, it is no trivial choice to focus on *or*. On the truth-conditional-cum-SI-view of *or* presented in the previous section, it is a key point of interaction between truth-conditional semantics and pragmatic inference. Logical disjunction is such a basic operator in logic and truth-conditional semantics, that it is fascinating to see that the interpretation of its natural language counterpart is so strongly determined by pragmatics. What could be a better place to look for regularities in pragmatics than in its interaction with a word that has such a basic meaning? Even its main competitor for the title of most studied SI-

trigger, *some*, does not have such a basic meaning, as it is subject to domain restriction and ambiguity between its quantifier, partitive and indefinite uses.

This thesis has three goals, which I present in this section. The first goal I already alluded to above: adding to a more structured theory of pragmatics by investigating the properties of the context that determine whether a scalar implicature arises. I discuss this goal in more detail in 3.1 below. In 3.2 I will explain the experimental approach and the methodological goal that follows from this, which is finding a suitable paradigm to test pragmatic inferences and SIs in specific. The third goal is assessing the psychological reality of SI-theory, and particularly of the SI-view of *or*. In my view the eventual goal of linguistics is modeling the language system of the actual language user. Therefore, we have to measure actual linguistic behavior in well-controlled experiments to test our theories. I discuss this goal in section 3.3.

3.1 The first goal of the thesis: investigating in which contexts SIs arise

In the previous section I pointed out that one of the hallmarks of conversational implicatures is that they can be cancelled. I illustrated this by showing that adding information to the contrary does not create a contradiction. However, it is widely acknowledged that SIs can also be absent if the context is such that they are not licensed by it. I use the expression 'be absent' here as some authors claim the SI is calculated but consequently cancelled implicitly in these cases, while others claim the SI is not calculated in the first place. I will not go into this discussion here (the interested reader is referred to section 4 of chapter 5).[5] What is important here is that whether a sentence gets interpreted with or without SI, depends on the wider linguistic context. Levinson (1983) formulates this idea as follows:

> '[...] implicatures can just disappear when it is clear from the context of utterance that such an inference could not have been intended as part of the utterance's full communicative import' (p. 115-116)
>
> 'So implicatures are defeasable and can drop out in certain linguistic or non-linguistic contexts.' (p. 116)

[5] In the rest of this work I will use sentences like 'The SI does not arise'. These should be taken theory-neutrally: they just means that at the end of the day, the SI is absent, and say nothing about whether the SI was calculated and cancelled or not calculated at all (unless explicitly stated otherwise).

In examples (17)-(19) in section 2 we saw how sentence-level properties affect SIs. When a scalar term is embedded under a Downward Entailing (DE) operator such as negation (*not*, *none*), the antecedent of a conditional (*if...*), or the restrictor of a universal quantifier (e.g. *Everyone who...*), the SI does not seem to arise. This shows that SIs are sensitive to structural properties of the sentences they appear in. Sentence-level properties that affect SIs are reasonably well-studied, also experimentally. Next to the behavior of scalar terms in sentences containing DE-operators (see e.g. Chierchia (2004) for theoretical work and Chierchia et al. (2001) for experimental work), the behavior of SIs in modal sentences and sentences with logical operators has been studied (see e.g. Geurts & Pouscoulous 2009 and Chemla & Spector 2010).

However, it is widely acknowledged that also in sentences in which no SI-blocking structural factors seem to be present, the SI is sometimes absent, due to the wider context that the sentence is part of. Very little is known about how properties of the wider linguistic context (anything bigger than the sentence-level) affect SIs. As Hirschberg (1985) puts it:

> 'While studies on conversational implicature have readily acknowledged the overriding importance of 'context' in the generation and interpretation of implicature, little attempt has been made to define how the particulars of an exchange interact with the conventions that underlie implicature.' (p. 140)

So although it is well-known that the wider context is important for the presence or absence of SIs, little attempt has been made in defining what the properties of the context are which determine whether SIs are present or absent. Usually, authors only mention the context-sensitivity of SIs without attempting to define what it is made up of. An example of this is the quote by Levinson (1983) above, where it is unclear what determines whether something is 'intended as part of the utterance's full communicative import'. It seems we are dealing with a wastebasket within the pragmatics wastebasket here. In order to explain the apparent ambiguity of *or* and other scalar terms, semantics moves a part of their meaning to pragmatics. Consequently, in order to keep up the SI view of e.g. *or* and explain its behavior on the sentence level, contextual effects are moved to the side. This dissertation is an attempt to take contextual properties that affect SIs out of the wastebasket of the pragmatics wastebasket.

Some theorists have explained context-sensitivity of SIs in terms of *relevance*, but defined this notion very broadly, which enables the theory to explain every example post hoc, but hardly predict anything. For instance

Relevance Theory (Sperber & Wilson 1986/1995) derives SIs by the *Presumption of Optimal Relevance*, given here:

> Presumption of Optimal Relevance (revised 1995 version)
> (a) The ostensive stimulus is relevant enough for it to be worth the addressee's effort to process it.
> (b) The ostensive stimulus is the most relevant one compatible with the communicator's abilities and preferences.

Carston (1998) shows how these principles explain presence or absence of SIs in certain contexts. However, the question when something is 'relevant enough' for a hearer, or what 'the most relevant utterance compatible with the communicator's abilities and preferences' is, seems to be hard to answer in such a way that it yields clear predictions. Relevance Theory describes an utterance as 'relevant enough' if it has 'a sufficient level of contextual effects' (e.g. leading to a conclusion about the world), which does not seem to get us any closer to a prediction. As Levinson (2000) puts it:

> '[...] I have suggested that SW-R [Sperber & Wilson Relevance Theory, AZ] is incapable of making clear predictions, partly because the theory is not clearly articulated but partly because the factor of cognitive effort, an essential ingredient in the proportional measurement of Relevance, is not empirically measurable (or at least not empirically measured).' (p. 57)

So Relevance Theory might be very good at explaining the presence or absence of SIs in individual cases after the fact, as far as I can tell it makes hardly any clear, testable predictions.

Fortunately, there are theories which do make testable predictions about a contextual property which determines whether an SI is calculated. These predictions will be tested in this dissertation. They are based on the property of *information focus*, which is a contextual property as it depends on material outside the sentence itself. Jackendoff (1972) proposed that a sentence can be split up into a part which contains new information, and a part which contains old information. The first part is called the *information focus* of the sentence, and the second part the *background*. This distinction can be clearly observed in question-answer sequences, such as (29):

(29) Q: What did Harry bring to the party?
A: Harry brought wine and bread$_F$ to the party.

In the answer in (29), *wine and bread* is the new information and therefore the information focus of the sentence. This is represented in (29) with the underlining and the *F* subscript. The rest of the sentence is the background. It contains information that was already present in the question, so it is old information.

Two theories which make predictions about the presence or absence of SIs based on this focus/background distinction are the topic-comment approach of Van Kuppevelt (1996) and the exhaustivity account of Van Rooij (2002). Van Kuppevelt proposes a framework in which discourse is organized as an ongoing questioning process. In his framework, the *comment* part of a sentence is the part that answers the contextual question. This corresponds to what Jackendoff calls the information focus, as is illustrated in (29) above. Van Kuppevelt claims SIs only arise if the constituent that the scalar term is part of appears in this part of the sentence. A similar prediction is made by Van Rooij (2002), but from a completely different perspective. He proposes an account in which SIs are derived by application of a covert exhaustivity operator. I will explain how this operator works in chapter 2. Crucially, the operator only exhaustifies the meaning of the material that is in the focus part. Hence, SIs are only predicted to arise in this part, which is the same prediction that Van Kuppevelt made.

What is interesting about this prediction for the current enterprise is that the information structure of the sentence, which part is focus and which part is background, depends on the wider context. It can for instance be determined by a contextual question to which the target sentence is an answer. Therefore, an account that SIs are focus-sensitive makes different predictions for the same sentence depending on the question it is an answer to. The focus-sensitive accounts thus pinpoint a contextual property which affects SI-calculation. Consider for instance the two question answer-pairs in (30) and (31):

(30) Q: What did Harry bring?
A: Harry brought <u>bread or chips</u>$_F$.

(31) Q: Who brought bread or chips?
A: <u>Harry</u>$_F$ brought bread or chips.

The same sentence is used as an answer in (30) and (31): *Harry brought bread or chips*. However, due to the different questions that were asked in (30) and (31), the focus structure of the answers is different. Only in (30) is the scalar term *or* part of the information focus of the sentence. The accounts of Van Rooij and Van Kuppevelt predict the SI of *or* should therefore arise in (30) and not in (31). This prediction will be tested in this dissertation.

In this thesis I take an experimental approach. In the next two subsections, in which the two other goals are introduced, I explain this approach.

3.2 The need for controlled experiments and the methodological goal

Above I argued it is fairly easy to come up with striking examples to illustrate a certain generalization, and that this is very common practice in pragmatics. Often general principles such as Grice's maxims or Relevance Theory's 'Presumption of Optimal Relevance' are argued for using examples which clearly trigger an intuition which can be explained by the principle. However, an inherent problem in pragmatics is that many different factors seem to be at work at the same time. As pragmatics is concerned with specific situations, speakers, contexts etc, there are numerous possible factors which might affect one particular intuition. Therefore, in order to find out what exactly made the example trigger the specific intuition, one has to isolate the relevant property and manipulate it systematically in well-controlled experiments. If minimal pairs can be created that only differ on the property under investigation, an experiment can confirm or disconfirm the hypothesis that this property is indeed responsible for the intuitions. Furthermore, parts of utterance meaning that are considered to belong to pragmatics are subject to a lot of individual variation, on the side of the contexts and expressions that give rise to them as well as on the side of the language users that interpret them. Therefore, we cannot rely on individual intuitions about specific contexts or expressions. In order to confirm or disconfirm hypotheses, we have to conduct experiments in which we gather data from a larger number of discourses and a larger number of language users. This is exactly the strategy I will follow in this investigation.

However, due to their context- and situation-dependent nature and their susceptibility to individual variation, pragmatic inferences are notoriously hard to assess in an experimental setting. Inferences that might arise in actual conversation can be hard to trigger in an experimental setting, and individual variation can blur the view of experimental results considerably. How hard to assess and how susceptible to individual variation pragmatic inferences are, is reflected in previous experimental research on SIs. Although as I said above SIs are considered to be one of the most robust inferences in pragmatics, a short but rich history of experimental research on SIs has shown that in general participants' behavior with respect to them is far less well-behaved than we would like. For example, in the landmark experiment by Noveck (2001), participants were asked to judge whether (French equivalents of) sentences of the form *Some elephants have trunks* were true

or false. Their behavior was far from univocal, as the percentage of 'false' answers due to the SI (*not all elephants have trunks*), was 59%. Geurts (in prep.) provides an overview of 11 well-known experiments on SIs, in which SI-rates vary between 25% and 65%, so none even coming close to 0% or 100%. The experimental research that has been done on the effect of structural factors on SIs also shows that the behavior of participants in experiments does not always match theoreticians' intuitions. Where there seems to be a consensus in the literature that in DE-environments SIs do not arise or are flipped (see 2.2 above), Chierchia et al. (in press) found around 50% SIs in two of these environments (see section 2.5 of chapter 3 for a more detailed discussion of their experiments). These examples show how tough it is to assess SIs experimentally.

So on the one hand we need well-controlled experiments in order to isolate the crucial factors in pragmatic inferences and rule out other factors, but on the other hand we want the experimental situation to be as natural an environment as possible for the inferences to arise. This is the methodological goal, or rather the methodological challenge, of the current work. Throughout this dissertation I will be in search of the best combination of test items and experimental paradigm to assess the hypothesis and SIs in general. I will start with the paradigm that is most often used to test SIs, the Truth Value Judgment Task (TVJT). After considering its shortcomings for testing SIs, I switch to a new experimental paradigm called the *Possible World Judgment Task* (PWJT). Finally, to avoid interference of an experimental question, I will turn to on-line experiments in which the Self-Paced Reading (SPR) paradigm is adopted.

However, the need for controlled experiments is not the only reason for taking an experimental approach. The goal of this thesis is not just to investigate the contextual properties which determine SI-calculation, but also to investigate the psychological reality of SI-theory in general, and in particular of the SI-view of *or*. I present this goal in the next subsection.

3.3 The psychological goal: the psychological reality of SI-theory

As I discussed in the previous section, the experimental approach allows us to test specific hypotheses from the theory, such as the dependency of SIs on information focus. However, it also allows us to investigate whether pragmatic theory, in this case SI-theory, is *psychologically real*. By this I mean whether it is a good model of the underlying knowledge of the language user and the actual processes of interpretation. For SI-theory, the question of psychological reality is whether it is a good description of the

representation of scalar terms in the mental lexicon of language users, and of the processes by which they interpret (sentences containing) scalar terms.

In this thesis I will test predictions stemming from SI-theory on the contexts in which SIs arise and in which they do not. Next to testing these predictions, the experiments contribute to determining whether SI-theory has psychological reality. If the pattern that was predicted by SI-theory is indeed observed in the experiments, this is an argument in favor of the view that the claims made by SI-theory about the underlying system are correct. Any competing account (for instance an ambiguity account of scalar terms), would have to explain how the observed behavioral pattern is predicted by that theory. Furthermore, in chapter 6 I present two reading time experiments which were designed to measure indications of SI-calculation during language processing, in the form of increased processing load. If this delay is only observed in conditions in which SI-theory predicts calculation of an SI, this is support for the psychological reality of SI-theory.

In this thesis I will particularly consider the psychological reality of the SI-view of *or*. The extension of Grice's theory made by Horn (1972) allows us to keep one meaning of *or* in the semantics, the (inclusive) logical disjunction meaning, avoiding a lexical ambiguity between inclusive and exclusive-*or*. I will investigate whether there are indications that the mental lexicon indeed only contains this one meaning of *or*, and that language users do indeed derive the exclusive meaning by calculating a scalar implicature. As we have seen above, from a truth-conditional semantics perspective it is an appealing move to make this assumption, and it is also in line with a conceptual argument against positing unnecessary ambiguities (Grice's 'Modified Occam's Razor'). However, in order to find out whether these theoretical arguments of SI-theory have any bearing on the language system of actual language users, we have to look at how language users actually behave when they interpret *or*. Again, the best way to study this actual behavior is in experiments in which we can control the conditions under which interpretation takes place.

I will asses the psychological reality of the SI-view of *or* by investigating whether its interpretation is focus-sensitive and whether we can find a reflex of SI-calculation when participants read sentences with *or*. However, I further scrutinize the psychological reality of the SI-view of *or* by investigating the relation between the interpretation of *or* and speaker expertise. It has often been noted that the standard Gricean derivation of SIs is problematic for *or*. As *or* typically gives rise to the inference that the speaker is not completely informed, this goes against one of the basic assumptions of the Gricean nonce derivation, exemplified in (26) above, the assumption that the speaker is well informed. Therefore, I will consider whether we can find experimental evidence for this problem of the SI-view

of *or*. I do this by asking participants about the knowledge state of the speaker and at the same time measuring their interpretation of *or*. As the nonce derivation is not the only proposal for how SIs are derived, I present a number of recent alternative theories of SI-calculation and consider whether they can account for this problem of speaker expertise. I will discuss more complex versions of the Gricean derivation (Sauerland (2004), Van Rooij & Schulz (2004)) as well as theories that claim SIs do not arise by Gricean reasoning, but in the grammar (Fox (2007) and Chierchia (2004, 2006)).

Finally, I also address the psychological reality of the SI-view of *or* by experimentally investigating the relation between the SI of *or* and exhaustivity. Many of the recent theories of how SIs are derived claim they come about by (some form of) the more general mechanism of exhaustification. This is the mechanism that gives rise to the intuition that the answer in (32) means that *only* John visited you:

(32) Q: Who visited you in the hospital?
A: John.

The claim that SIs and exhaustivity are the result of one mechanism gives rise to an interesting prediction. This is the prediction that when sentences like (33) give rise to the SI in (34), they should also give rise to the exhaustivity inference in (35):

(33) Harry brought bread or chips.

(34) Harry did not bring bread and chips.

(35) Harry did not bring something else (besides bread or chips).

I argue this prediction follows from all of the recent accounts mentioned above, and I test it by measuring both types of inference in one experiment.

Summing up, in this thesis I present four types of data which address the psychological reality of the SI-view of *or*: the effect of focus on the SI of *or*, the real-time processing of sentences with *or*, the relation between speaker expertise and the SI of *or*, and the co-occurrence of the SI of *or* and exhaustivity.

A well-known criticism of the experimental approach in pragmatics is that some pragmatic theories were not designed to be actual theories of linguistic competence or behavior. For instance, it is often claimed that Grice's theory was not meant to explain or predict certain behavior, but that it is a *normative* theory: a theory about what speakers and hearers are allowed to do, or what a certain expression in a certain environment ought to

imply. As such, it does not make any predictions that language users will actually behave according to these norms. However, as Geurts (in prep.) points out, it is a small step from Grice's theory to predictions about competence and actual behavior. Moreover, it seems to be a fairly successful move, as it can explain a wide range of empirical facts. Geurts also argues convincingly against objections that Grice's theory is psychologically implausible. Moreover, despite that the theory might not be designed to account for actual behavior, it is the dominant view in the literature about the interpretation of *or*. Therefore, I will ignore objections of this sort, as many authors have done before me, and consider all theories presented in this dissertation, including Grice's, as making predictions about actual linguistic behavior. As Katsos (2006) nicely puts it:

> '[...] distinctions between actual inferences and normative implications are at best an idle philosophical point for the linguist that subscribes to a cognitive view of semantics and pragmatics.' (p. 75)

Finally, the experimental approach has an additional goal that goes beyond the specific hypotheses or theories that the experiments are set up to test. This is the goal of measuring linguistic behavior by itself, i.e. finding out what people actually do in certain conditions. The data can expose other regularities than the ones predicted by the theory, which can provide inspiration for new theories. Also, if new theories take the place of the old ones for other reasons, the experimental data do not get thrown away together with the theories. As they are still valid measurements of actual linguistic behavior, the new theories will have to account for them too. For this reason, adding to the pool of experimental data contributes to the progress in linguistics irrespective of the particular theory under investigation. As long as this pool keeps growing, new theories of linguistic competence will have to have ever increasing explanatory adequacy, which means they have to be better theories. Therefore, an additional goal of this thesis is to provide experimental data on how language users interpret scalar terms, and *or* in particular. I will now provide the outline of the thesis and preview the most important experimental results.

4. Outline of the thesis and preview of the results

In chapter 2 I present the theoretical background of the predictions based on information focus. As the notion *focus* is used in many different ways in the literature, I will first introduce the notion as it will be used in this work, based on the focus-background distinction by Jackendoff (1972). Then I

shortly introduce some more recent theories of the semantics of focus. There is a similarity between focus and SIs in the sense that both rely on alternatives, which has motivated some authors to extend the semantics of focus to SIs (Krifka (1995), Chierchia (2004)). I briefly discuss these proposals in chapter 2. However, these accounts are mainly concerned with the sentence level and so-called 'neutral' contexts and hence they make no claims about contextual properties that trigger or block SIs. Finally, I will introduce the two theories which provide the hypothesis of focus-sensitivity of SIs, the accounts of Van Kuppevelt (1996) and Van Rooij (2002). This hypothesis will be tested experimentally in chapters 3-7.

In chapter 3 I present three experiments in which the Truth Value Judgment Task (TVJT) paradigm was used to test the focus-sensitivity of SIs, specifically the SI of *or*. In these experiments participants were asked to judge whether a target sentence containing *A or B* is true relative to a story in which both A and B are the case. A 'false' answer indicated that they calculated the SI. In Experiments 1 and 2, in which a written version of the TVJT was used, the target sentence was an answer to an explicit question. The focus structure of the target sentence was therefore manipulated by changing this question between conditions, as in (30) and (31) above, repeated here.

(30) Q: What did Harry bring?
A: Harry brought bread or chips$_F$.

(31) Q: Who brought bread or chips?
A: Harry$_F$ brought bread or chips.

The results of these experiments showed that participants were indeed more likely to calculate the SI if the scalar term is part of the information focus. In Experiment 3 I used spoken stimuli in which information focus was signaled by stress. Due to this setup, the explicit questions could be left out, ruling out interference of repeated material from the question. Again, more exclusive interpretations of *or* were observed when it was part of the information focus.

In chapter 4 I present Experiments 4-7, in which I used a new experimental paradigm called the *Possible World Judgment Task* (PWJT) to assess the effect of focus on SIs. In the PWJT, participants were again presented with stories and target sentences with *A or B*. However, contrary to the TVJT, the actual situation was left out and participants were asked whether they considered the *A and B* situation to be possible. A negative answer indicated they calculated the SI for the target sentence. As the participants did not know the actual situation, this task provides a more

natural environment for SIs to arise than the TVJT. Next to avoiding the shortcomings of the TVJT for testing SIs (discussed at the end of chapter 3), the goal of Experiments 4-7 was to tease apart focus and another contextual property: the relevance of the stronger scalar alternative (*alternative relevance* from now on). One of the necessary conditions for SI-calculation by the nonce derivation discussed above, is that the stronger scalar alternative is relevant to the hearer. Whether this was the case was left unspecified in the experimental items of Experiments 1-3. One could argue that the difference between the questions, which were of the type of (30) and (31) above, also introduced a difference in alternative relevance between the two conditions. It could be argued that questions like (30) indicate the scalar alternative *Harry brought bread and chips* was relevant to the questioner, while questions like (31) do not. This difference could be responsible for the difference in SI-rates, instead of focus. To tease apart focus and alternative relevance, I explicitly manipulated both focus and alternative relevance in the conditions of Experiments 4-7. In two of four experiments of Experiments 4-7 an effect of focus and alternative relevance together was observed. However, the differences between the conditions in which the two properties were contrasted were too small to draw firm conclusions about the relative importance of the properties separately. The small differences were probably a side-effect of the experimental paradigm, in which the experimental question overruled the critical manipulations. Therefore, I turned to on-line (processing) experiments, in which the explicit question could be left out.

Before I present the processing experiments in chapter 6, I discuss some previous processing experiments on SIs from the literature in chapter 5, in order to find out what the on-line reflex of SI-calculation might look like. Most of these experiments were designed to settle two heated debates on SIs: the globalist-localist debate and the defaultist-contextualist debate, both of which I address briefly. I present a number of processing experiments which tried to settle these debates. I consider the self-paced reading experiments of Katsos (2006) in detail, as these form the basis for the reading time experiments in chapter 6. Katsos and colleagues found that (segments containing) scalar terms are read slower in (intuitive) SI-triggering contexts than in (intuitive) SI-blocking contexts, suggesting SI-calculation leads to increased processing cost. They also found that reading of material in the next sentence which is in line with an SI is facilitated in SI-triggering contexts. I take their prediction of a delay due to SI-calculation as the crucial prediction for the on-line experiments. If more SIs are calculated in the focus-condition, this should be reflected by longer reading times on the scalar region.

Experiments 8 and 9, presented in chapter 6, tested this prediction with self-paced reading (SPR), an experimental paradigm in which the participant reads trough a text piece by piece, controlling the transition from one piece to the next herself. Again, contexts containing question-answer pairs such as (30) and (31) were used, and reading times were measured on the *A or B* region of the answer. I also included control conditions with *and* instead of *or*, which controlled for other possible differences between the conditions. In Experiment 8, in which only focus was manipulated, no delay was observed on the scalar region (after controlling for a repetition effect). This could be due to the fact that as the context was very minimal and nothing hinged on the SI, there was no incentive for participants to calculate it. Therefore, I used richer contexts in Experiment 9, also manipulating alternative relevance. In this experiment I also included another measure: a following region in which either a singular pronoun (*it*) or a plural pronoun (*them*) was used to refer back to *A or B*. As a plural pronoun is incompatible with an exclusive reading of *or* (a reading with SI), this should lead to a delay in reading. The results confirmed the prediction that in the focus condition there was both a delay on the scalar region itself and on the plural pronoun. These processing delays were not observed for the non-focus conditions, the *and* control conditions and the conditions with a singular pronoun continuation. This supports the hypothesis that more SIs were calculated when *or* was in the information focus part of the sentence.

In chapter 7 I extend the scope of the investigation and present two TVJT experiments that tested whether the results of focus-sensitivity of SIs can be replicated with another scalar term: *most* (which gives rise to the SI *not all*) in another structural position (subject position instead of direct object). I also test the predictions of the focus-sensitive view for answers to yes/no-questions. Finally, I explore the interpretation of scalar terms in yes/no-questions themselves. Although the differences were smaller than with *or*, more SIs were calculated when *most* was part of the focus, suggesting that the generalization indeed extends to a wider range of scalar terms and conditions.

In chapter 8 I address the relation between the SI of *or* and speaker expertise. The use of *or* normally gives rise to the inference that the speaker is not completely informed. This is problematic for the Gricean SI-view of *or*, as the nonce derivation crucially relies on the assumption that the speaker is well informed. I present this problem in more detail and I call it the *Speaker Expertise Paradox*. I also present additional results from Experiment 4, in which all test items next to an SI-question also contained a question about speaker expertise. In this question, the participant was asked whether she considered the speaker in the story to be fully informed. The results show that participants often indeed interpreted *or* as exclusive-*or*

while they did not consider the speaker to be fully informed, which goes against the Gricean view. In the second part of chapter 8 I consider four alternative theories, the two Gricean alternatives by Sauerland (2004) and Van Rooij & Schulz (2004) and the non-Gricean alternatives of Fox (2007) and Chierchia (2004, 2006) / Chierchia et al. (2008). I argue all four do not solve the Speaker Expertise Paradox.

The relation between the SI of *or* and exhaustivity is considered in chapter 9. The theory of Van Rooij (2002), on which the prediction of focus-sensitivity was based, considers SIs to be a form of exhaustivity inference. This gives rise to the interesting prediction exemplified in (33)-(35) above, that if *A or B* gives rise to the SI *not A and B*, it should also give rise to the exhaustivity inference that *not C*. I show not only Van Rooij (2002)/Van Rooij & Schulz (2004) make this prediction, but also the grammatical exhaustification theory of Chierchia et. al (2008). I present additional data from Experiments 5-7 which address this prediction. Next to the SI-questions, Experiments 5-7 also contained items that asked whether the participants considered it possible that other (non-scalar) alternatives held. The results show that in the same conditions, the SI-rates (the percentage exclusive-*or* interpretations) and the external exhaustivity-rates are not always in the same range. Exclusive-*or* is more frequent that external exhaustivity, contrary to the predictions by the accounts mentioned above.

In chapter 10 I summarize the thesis and I draw conclusions on the three goals presented in section 3 above. The main conclusions are that the focus-sensitivity of SIs was confirmed in a wide range of experimental circumstances, but that none of the current SI-theories can explain the full range of data presented in this thesis on the interpretation of *or*, which casts considerable doubt on the psychological reality of the SI-view of *or*.

5. Summary of the research program

This dissertation is aimed at contributing to a more structured theory of pragmatic inference and investigating the psychological reality of these inferences. To achieve this goal I focus on one of the most robust and widespread pragmatic phenomena: scalar implicature. Particularly, I look in detail at the SI of *or*. This lexical item is very suitable for such an investigation, as (in the relevant cases) it is claimed to have a purely truth-conditional basic meaning (logical disjunction), but its interpretation in natural language is strongly affected by pragmatics. Therefore, it is the ultimate case of where semantics and pragmatics meet.

This thesis addresses the question when SIs arise. There seems to be a gap in the study of SIs as to what exactly the properties of the context are

that trigger or block them, irrespective of a general consensus in the literature that their presence depends crucially on context. This study contributes to the filling of this gap, specifically by considering the property of information focus, as determined by an explicit or implicit contextual question. The main hypothesis of this part of the investigation is that SIs only arise if the scalar term is in the part of the sentence that forms the information focus. This hypothesis is confirmed in this thesis in a wide range of experimental circumstances.

As pragmatic inferences are subject to a lot of individual variation, both between the linguistic material that gives rise to them and between the language users that draw them, the best way to assess them is by conducting controlled experiments. In these experiments the property under investigation can be isolated and its effects can be observed over larger groups of items and speakers. I adopt this experimental strategy in this thesis. However, as SIs are notoriously hard to assess in an experimental setting, the methodological challenge of the thesis is to find the right paradigm for doing so. Therefore, in this work a series of 11 experiments is presented, set up in three different experimental paradigms, two of which (the Truth Value Judgment Task and the new Possible World Judgment Task) employ off-line judgments and one of which (Self-Paced Reading) measures processing through reading times.

The experimental approach also allows us to investigate the psychological reality of SI-theory in general, and specifically the SI-view of *or*. This thesis addresses this question by presenting experimental data on the effect of focus on the SI of *or*, the real-time processing of sentences with *or*, the relation between speaker expertise and the SI of *or*, and the co-occurrence of the SI of *or* and exhaustivity. Taken together, the data show that the SI-view of the meaning of *or*, which is so often taken for granted, is not so obvious at all.

CHAPTER 2: SCALAR IMPLICATURES AND FOCUS

1. Introduction

As I pointed out in the previous chapter, very little is known about the properties of the context that are responsible for triggering or blocking scalar implicatures (SIs), although the role of context in the generation of SIs is widely acknowledged. In this chapter I will consider two theories that make clear predictions on how the context determines whether or not an SI arises, based on the notion of *information focus*. Although this is a property of a part of a sentence, we will see that it is determined by the wider context. Therefore, it makes different predictions for one and the same sentence dependent on the context it appears in.

In section 2 I provide some theoretical background on the semantics of focus. The notion of information focus dates back to at least Jackendoff (1972), a work which served as a starting point for much later work on the semantics of focus. I discuss his approach in section 2.1, followed by a short introduction to more recent theories of the semantics of focus, based on a phenomenon called *Association with Focus*. The similarities between focus and scalar implicatures have led some authors to propose theories of SIs based on focus semantics. Two of these, Krifka (1995) and Chierchia (2004)/(2006), will be discussed in section 3. However, these theories are mainly concerned with the sentence-level and try to account for SIs in so-called *neutral* contexts, so they make no clear predictions about properties of the wider context that give rise to or block SIs. The two theories of SIs that do make such predictions, the topic/comment approach of Van Kuppevelt (1996) and the exhaustivity view of Van Rooij (2002), which provide the hypothesis that will be tested in this thesis, are discussed in section 4. Both theories are not directly derived from focus semantics in the way the accounts of Krifka and Chierchia are, but they make predictions about the contexts in which SIs arise based on the focus-background distinction.

2. The semantics of focus

2.1 The focus-background distinction: Jackendoff (1972)

One of the first accounts of the semantics of focus is Jackendoff (1972). He provides the example dialogue in (1), where capitals indicate the main stress of the sentence.

(1) Q: Is it JOHN who writes poetry?
A: No, it is BILL who writes poetry.

He argues the answer in (1) can be split up in two parts: the new information *Bill*, and the old information, the part that is repeated from the question (the rest of the sentence). He calls these parts respectively the *focus* and the *presupposition*, and proposes the following working definitions:

> 'As working definitions, we will use "focus of a sentence" to denote the information in the sentence that is assumed by the speaker not to be shared by him and the hearer, and "presupposition of a sentence" to denote the information in the sentence that is assumed by the speaker to be shared by him and the hearer.' [p. 2 of ch. 6]

In (1), the information *Bill* is not assumed by the answerer to be shared by the questioner, so it is the focus, while the information that someone writes poetry is assumed to be shared, so it is the presupposition. Nowadays, Jackendoff's presupposition (the old information) is usually called the *background*, and *presupposition* has a different meaning in modern pragmatics (see section 1 of chapter 1). Jackendoff states that this division of the meaning of a sentence into a focus and a presupposition/background part is an aspect of the semantic representation (SR) of the sentence. He furthermore argues that this division is somehow derived from the syntax, by a rule which he calls *focus assignment*. This rule adds the focus-background division to the semantic representation.

Other theories of those days (e.g. Chomsky 1971) derived the focus-background distinction from the stress structure of the sentence, claiming the focus of the sentence is a phrase containing the main stress. Jackendoff however points out that this is not right. Consider (2):

(2) Q: Did Fred HIT Bill?
A: No, he KISSED him.

In the answer in (2), the V *kissed* is stressed. However, the VP *kissed him*, although it contains the main stress, cannot be the focus, as *him* (Bill) is old information. Jackendoff concludes that 'containing the main stress is a necessary but not sufficient condition for a phrase to be focus' (p. 9 of ch. 6). Crucially, he turns the dependence of stress and focus around, by stating stress depends on focus:

> 'If a phrase P is chosen as the focus of a sentence S, the highest stress in S will be on the syllable of P that is assigned highest stress by the regular stress rules.' (p. 9 of ch. 6)

Jackendoff catches the generalization that focus and stress are related by arguing that stress assignment is one of the results of the fact that a constituent is marked as focus in the syntactic structure. Therefore, he assumes a syntactic marker F, which can attach to any node in the syntactic representation. This syntactic marker has effects on two systems: the semantic system, giving rise to the focus-background division, and the phonetic system, triggering the main stress of the sentence to be on (a syllable of) the focus.

Jackendoff proposes a three step derivation for how the rule of focus assignment adds the focus-background distinction to the semantic representation , and illustrates this with example (3):

(3) John LIKES Bill.

The first step is that focus assignment splits the semantic representation into two formal objects. The first object is the meaning of the F-marked nodes, called *Focus* (with a capital *F*), and the second object is a one-place predicate *Presupp(x)*, which is derived by replacing the Focus by an appropriate semantic variable *x*.[1] Jackendoff describes the restrictions on this variable as follows:

> '[...] the variable must be chosen in such a way that it defines a coherent class of possible contrasts with the focus, pieces of semantic information that could equally well have taken place of the focus in the sentence, within bounds established by the language, the discourse, and the external situation.' (p. 15 of ch. 6)

For (3), Jackendoff says Presupp(*x*) is something like (4):

(4) the {relation between John and Bill}
{attitude of John toward Bill} is *x*

So the 'coherent class of possible contrasts with the focus' is the sum of the class of relations between John and Bill and the class of attitudes of John towards Bill. The second step of the focus assignment rule is constructing a presuppositional set, the set of values which can be substituted for *x* in

[1] I guess the capitals are used to indicate that these are formal objects.

Presupp(x), and yield a true proposition.[2] For (3), these are the other relations between John and Bill or other attitudes of John towards Bill that are true. Jackendoff represents this presuppositional set as λxPresupp(x).

In the third step the presupposition and the assertion are constructed from the presuppositional set. The presupposition is simply that the presuppositional set is under discussion (it is what we are talking about), so it has a form like (5):

(5) Presupposition: λxPresupp(x) {is a coherent set
{is well-defined
{is amenable to discussion
{is under discussion

The assertion is that the focus is a member of the presuppositional set, see (6):

(6) Assertion: Focus ∈ λxPresupp(x)

For the example in (3) the presupposition is (7) below and its intuitive paraphrase is (8). The assertion is (9), paraphrased in (10):

(7) λx[the {relation between John and Bill} is x] is {well defined
{under discussion

(8) We're talking about possible {relations between John and Bill.
{attitudes of John toward Bill.

(9) like ∈ λx [the {relation between John and Bill} is x]
{attitude of John toward Bill}

(10) *like* is one of {the relations between John and Bill.
{John's attitudes toward Bill.

In the literature many alternative proposals have been made for the way the assertion and the presupposition/background are derived from the syntactic structure with F-marking (see the next section for examples). However, Jackendoff's idea that focus-marking is a grammatical device that has different reflections in the different representations (phonological representations and semantic representations), is widely adopted. Also, the

[2] This is where Jackendoff's analysis is crucially different from e.g. Alternative Semantics, in which there is no constraint of truth of alternatives.

distinction between focus and background has been applied to many linguistic phenomena, scalar implicatures being one of them. But before I get to that, I will discuss an important phenomenon in the theory of focus. This will give me the opportunity to introduce some more recent theories of the semantics of focus, as Jackendoff acknowledged that for this particular phenomenon 'at the present state of research the formal nature of this rule cannot be specified' (p. 21 of ch. 6). The phenomenon is called *Association with Focus*.

2.2 Association with Focus: Structured Meanings and Alternative Semantics

Association with focus is the phenomenon that certain natural language operators, like *only* and *even*, contribute meaning to the sentence that depends on the focus of the sentence. The discussion in this subsection is based on Krifka (2006), who provides the following examples of Association with Focus:

(11a) John only introduced $\underline{\text{Bill}}_F$ to Sue.

(11b) John only introduced Bill to $\underline{\text{Sue}}_F$.

(11a) means that the only person John introduced to Sue is Bill, while (11b) means that Sue is the only person that John introduced Bill to. Krifka concludes that in order to account for the difference in meaning between (11a) and (11b) in compositional semantics, we have to assume that the meaning of the two VPs is different. There are two frameworks which account for this type of focus data, the *Structured Meanings* (SM) approach (Jacobs 1983, Von Stechow 1990) and the *Alternative Semantics* (AS) approach (Rooth 1985, 1992).

On the SM approach, the meaning of the VP in (11a) and (11b) is structured, in the sense that it is ordered in focus (F), alternatives (A), and background (B), in an ordered tuple <F,A,B>. The meaning of the VP in (11a) is (12):

(12) $[\![\text{introduced } \underline{\text{Bill}}_F \text{ to Sue}]\!] = \langle \text{BILL}, A, \lambda x[\text{INTROD}(\text{SUE})(x)] \rangle$

So in (12), BILL is the focus, A is the set of alternatives to *Bill* and $\lambda x[\text{INTROD}(\text{SUE})(x)]$ is the background. A contains *Bill* and at least one other element. This set of alternatives is contextually determined, so for (12) it will contain other individuals in the domain. It is comparable to Jackendoff's

presuppositional set, with the difference that the elements do not have to yield a true proposition. Focus-sensitive operators such as *only* take structured meanings like (12) and turn them in to regular meanings. For instance, the meaning of *only* is (13):

(13) $\text{ONLY}(<F,A,B>) = \lambda x \forall Y \in A[B(Y)(x) \rightarrow F=Y]$

Now if we apply this meaning of *only* to (12), we get (14) as the meaning of the VP, and (15) as the meaning of the sentence:

(14) $\text{ONLY}(<\text{BILL}, A, \lambda x[\text{INTROD}(\text{SUE})(x)]>$
$= \lambda x \forall y \in A[\text{INTROD}(\text{SUE})(y)(x) \rightarrow y = \text{BILL}]$

(15) $\forall y \in A[\text{INTROD}(\text{SUE})(y)(\text{JOHN}) \rightarrow y = \text{BILL}]$

This sentence meaning is that every person that John introduced to Sue is Bill, which is intuitively right.

The challenge for the SM approach is how *only* has access to the meaning of the focus (in other words, how the focus-background structure is represented). One option is that the structure is reflected by movement at LF. The LF for (11a) would then look something like (16), with the focus moved up to a higher position:

(16) [Bill$_i$ [introduced t_i to Sue]

On the AS approach (Rooth 1985, 1992) expressions in focus cannot be accessed directly. The meaning is derived from the meaning of the whole scope of the focus-sensitive operator (e.g. in our example the scope of *only*, so the VP), and the alternatives introduced by the focus element. The idea is that every expression has two meanings: its ordinary meaning $[\![...]\!]$, and a *focus semantic value*, $[\![...]\!]^f$ which is the set of propositions formed by substituting the focus with contextually given alternatives. When function application takes place, the focus semantic values are also combined by function application, generating a new set of alternatives. This is formalized in (17):

(17) If $[\![(a\ b)]\!] = f([\![a]\!], [\![b]\!])$
then $[\![(a\ b)]\!]^f = \{f(X,Y) \mid X \in [\![a]\!]^f, Y \in [\![b]\!]^f\}$

Crucially, the set of alternatives of a non-focus expression *a* is a singleton set of its ordinary meaning: $\{[\![a]\!]\}$, while the set of a focus expression is some non-singleton set $\text{ALT}([\![a]\!])$, containing the element itself and

contextually given alternatives. Applied to our example (11a), we get the ordinary meaning of the VP in (18), and the alternative set in (19):

(18) $[\![\text{introduced } \underline{\text{Bill}}_F \text{ to Sue}]\!] = \lambda x[\text{INTROD}(\text{SUE})(\text{BILL})(x)]$

(19) $[\![\text{introduced } \underline{\text{Bill}}_F \text{ to Sue}]\!]^f =$
$\{\lambda x[\text{INTROD}(\text{SUE})(y)(x) \mid y \in \text{ALT}(\text{BILL})\}$

In AS, the meaning of *only* is the following:

(20) $[\![\,[\text{only}[_{VP}\ \alpha]]\,]\!] = \lambda x[[\![\alpha]\!](x) \wedge \forall Y \in [\![\alpha]\!]^f[Y(x) \rightarrow Y = [\![\alpha]\!]\,]$

If we apply this to the VP we get (21) as the meaning of the VP, and (22) as the meaning of the sentence.

(21) $\lambda x[\text{INTROD}(\text{SUE})(\text{BILL})(x)] \wedge \forall Y \in \{\lambda x[\text{INTROD}(\text{SUE})(y)(x) \mid$
$y \in \text{ALT}(\text{BILL})\}[Y(x) \rightarrow Y = \lambda x[\text{INTROD}(\text{SUE})(\text{BILL})(x)]$

(22) $\text{INTROD}(\text{SUE})(\text{BILL})(\text{JOHN}) \wedge \forall Y \in \{\lambda x[\text{INTROD}(\text{SUE})(y)(x) \mid$
$y \in \text{ALT}(\text{BILL})\}[Y(\text{JOHN}) \rightarrow Y = \lambda x[\text{INTROD}(\text{SUE})(\text{BILL})(x)]$

The sentence meaning in (22) is according to Krifka: '[...] it is claimed that John introduced Bill to Sue, and for all properties of the type 'introduce y to Sue', where y is an alternative to Bill, if John has property Y, then Y is the property 'introduce Bill to Sue.' (p. 5), which is indeed the intuitive reading of (11a).

As in this approach the meaning of *only* does not rely on the meaning of the focus expression itself, it does not have the problem that SM has about how this access comes about. Rather, it uses the distinction between alternative sets of focus and non-focus expressions: as only alternative sets of focus elements contain more than its ordinary meaning, they are the only elements that can contribute to the compound focus alternative sets.

There is a close similarity between focus semantics and scalar implicatures, in the sense that both rely on the presence of alternatives. Therefore, it is not surprising that views from focus semantics have been extended to account for SIs. I will discuss two examples in the next section.

3. Scalar Implicature Theory based on Focus Semantics

3.1 Krifka (1995)

One of the proposals in which focus theory was extended to account for SIs, is the proposal by Krifka (1995). He derives SIs by introducing an assertion operator *Assert*, which applies to Structured Meanings, <B,F,A> triples, discussed in the previous section. He assumes a common ground *c* and common ground update by propositions (c∩p) along the lines of Stalnaker (1972). These are given in (23) and (24):

(23) The participants of a conversation assume, for every stage of the conversation, a mutually known common ground c.
(24) If one participant asserts proposition p, and the audience does not object, the common ground c is restricted to c∩p.

Krifka introduces the operator *Assert*, which applies to a <B,F,A> triple, and updates the common ground with the proposition B(F), with a number of restrictions. See (25):

(25) Assert(<B, F, A>)(c) = c ∩ B(F), iff
   a) B(F) is assertable (p expresses something that isn't already established and isn't taken to be impossible).
   b) For all F'⊆ A such that c ∩ B(F') ≠ c ∩ B(F): the speaker has reasons not to assert B(F').
   c) There are F'∈ A such that B(F') is assertable w.r.t. c and c∩B(F') ≠ c∩B(F).

The important condition for SIs is the (b) condition. As Krifka points out, for a proposition in which the focus F is a scalar term, the proposition asserted and its alternative propositions (which are formed by replacing F with other members of A), are ordered by semantic strength (through entailment). Consider for instance sentence (26), where *most* is the focus.

(26) John ate $\underline{\text{most}}_F$ of the cookies.

The set of alternatives A in this case is a set of scalar quantifiers, e.g. the set {some, most, all}. So any alternative proposition formed by replacing F (*most*) by an F', which is one of the alternatives, will either entail the assertion, or be entailed by it. For instance, if we take F' to be *all*, B(F') entails B(F), and if we take F' to be *some*, B(F) entails B(F'). Krifka argues that in these cases, if the speaker wants to be truthful and informative, there

are two reasons for not asserting an alternative proposition: either it leads to a weaker assertion, or the speaker considers it to be false. The first is the case if the alternative assertion B(F') is entailed by B(F). In our example, if we take F' to be the weaker *some*, B(F') would be a weaker statement than B(F), and that is the reason why the speaker chose not to utter B(F'). The second reason, that the speaker considers the alternative to be false, is the case if the alternative B(F') entails B(F). If we take F' to be the stronger *all* in our example, B(F') would entail B(F), so it would be stronger. Therefore, it has to be the case that the speaker considers this alternative to be false. This way the SI *it is not the case that John ate all of the cookies* is derived for (26).

Krifka formalizes this rule by which SIs are derived by introducing another operator, *Scal.Assert*. This is a special case of the general assertion rule, reserved for cases in which the alternative assertions are ordered by semantic strength, so the cases in which a scalar term is the focus. This condition is formalized in (27):

(27) Assert(<B,F,A>)(c) = Scal.Assert(<B,F,A>)(c), if for all F'$\in$ A:
$[c \cap B(F')] \subseteq [c \cap B(F)]$ or $[c \cap B(F)] \subseteq [c \cap B(F')]$

The rule for the SIs is then formalized as in (28):

(28) Scal.Assert(<B,F,A>)(c) =
$\{i \in c \mid i \in B(F) \wedge \neg\exists F' \in A[([c \cap B(F')] \subset [c \cap B(F)]) \wedge i \in B(F')]\}$

This rule says that the common ground is only updated with the worlds i in which B(F) is true, and in which there is no alternative assertion with Focus F' which would entail B(F). I now turn to a more elaborate account of SIs based on focus semantics: the proposal by Chierchia (2004) and (2006).

3.2 Chierchia (2004) and (2006)

Chierchia (2004) extends the mechanism for focus of Alternative Semantics to scalar terms. As in Alternative Semantics, he claims that every expression has a plain meaning $[\![\alpha]\!]$, and a meaning which is based on a set of alternatives $[\![\alpha]\!]^{ALT}$. However, for scalar terms the set of alternatives comes about by substituting the scalar term with other items of the scale. The set of alternatives is therefore defined as (29):

(29) $[\![\alpha]\!]^{ALT}$ = { $\{a_1,..., a_n\}$, if α is part of a scale $<a_1,... ,a_n>$
{a} otherwise

The alternatives on the scale rely on a lexically given scale (e.g. a Horn scale). In Alternative Semantics for focus, the focus semantic value $[\![\ ]\!]^f$ of a non-focus expression *a* was the singleton set {a}, containing only the plain meaning, and the focus semantic value of a focus expression contained the contextually given alternatives. Similarly, if in Chierchia's system an expression does not contain a scalar term, $[\![\alpha]\!]^{ALT}$ is the set containing only the plain meaning, and if it does contain a scalar term, $[\![\alpha]\!]^{ALT}$ contains the scalar alternatives.

The meaning which is based on the set of alternatives, is called the 'scalar' or 'strengthened' meaning, represented by $[\![\alpha]\!]^S$. This strengthened value comes about by rule (30), which Chierchia himself calls a version of Krifka's rule, given in (28) above.

(30) If φ is a scope site (of type t), then $[\![\varphi]\!]^S = [\![\varphi]\!]^S \wedge \neg S([\![\varphi]\!]^{ALT})$

In this formula, $S([\![\varphi]\!]^{ALT})$ is the member of the alternative set immediately stronger than φ (so the expression with the scalar term replaced by a scalar term that is one step higher on the scale). This rule does indeed return the same results as Krifka's rule (28). While Krifka's (28) says there is no stronger (entailing) scalar alternative which is true, Chierchia's (30) explicitly states the stronger alternative is negated.

Similar to function application with focus values in Alternative Semantics (see (17) above), on Chierchia's approach function application takes place with the strengthened meanings. However, function application of strengthened meanings does not always lead to a stronger result than if the plain meaning is used. This is for instance the case in DE-environments, where the strength of the alternatives is reversed (e.g. *not or* entails *not and*). So function application of strengthened meanings would lead to wrong predictions in those environments. Therefore, Chierchia introduces the Strength Condition:

(31) Strength Condition: The strong value cannot become weaker than the plain value.

So in DE-environments, instead of using the strengthened meaning as an input to function application, the plain meaning should be used. Chierchia proposes the following rule for function application of constituents α and β, which he calls *Strong Application*:

(32) $[\![\alpha\ \beta]\!]^S = \{$ $\{[\![\alpha]\!]^S([\![\beta]\!]^S)$ if $[\![\alpha]\!]$ is not DE.
$\{[\![\alpha]\!]^S([\![\beta]\!])$ otherwise

So crucially, in DE-contexts, the plain meaning of $[\![\beta]\!]$ enters function application, instead of its strengthened meaning $[\![\beta]\!]^S$, as the plain meaning gives rise to a stronger meaning. For example, if α is negation and β is *has A or B*, we should select the plain meaning of *A or B* ($A \vee B$) instead of the strengthened meaning ($(A \vee B) \wedge \neg(A \wedge B)$), as the plain meaning gives a stronger result.

However, Chierchia points out that in DE-environments new SIs can arise, e.g. the SI from *not A and B* to *A or B*. To account for this, we have to consider the result of function application of $[\![\alpha]\!]$ and the alternatives of $[\![\beta]\!]$. While in a UE-environment *and* is the top item of the scale, in a DE-environment it is weaker than *or*. Therefore, we have to take into account the result of function application of $[\![\alpha]\!]$ and the alternative of β with *or*. Chierchia proposes to add this to the definition of Strong Application, see (33):

(33) $[\![\alpha\ \beta]\!]^S = \{$ $\{[\![\alpha]\!]^S([\![\beta]\!]^S)$ if $[\![\alpha]\!]$ is not DE.
$\{[\![\alpha]\!]^S([\![\beta]\!]) \wedge \neg S([\![\alpha]\!]([\![\beta]\!]^{ALT}))$ otherwise

So if we apply function application of a DE α and another expression β, we take the plain meaning of β, but we also rule out stronger results of application of α and alternatives to β, which accounts for what Chierchia calls 'indirect implicatures': implicatures that are the result of the reversal of strength of scalar alternatives in DE-environments.

In Chierchia (2006), the exhaustivity operator O is introduced, to replace the rule in (30) above. Chierchia defines O as in (34), where C is the set of alternatives to an expression φ, $[\![\varphi]\!]^{ALT}$.[3]

(34) $O_C[q] = q \wedge \forall p[((p \in C) \wedge p) \rightarrow (q \subseteq p)]$

The strengthened meaning of an expression φ, represented by $[\![\varphi]\!]_S$, is the result of applying O to the plain meaning and its alternatives, see (35):[4]

(35) $[\![\varphi]\!]_S = O_C\,[\![\varphi]\!]$, where $C = [\![\varphi]\!]^{ALT}$

The mechanism of O is similar to the mechanism of Chierchia (2004) discussed above, and Krifka's rule in (28) above. The operator applies to a meaning q and says that every alternative p from the set of alternatives C is

[3] I added some brackets to the formula in (34) for disambiguation

[4] For some reason, the superscript S in Chierchia (2004) is replaced by a subscript S in Chierchia (2006).

entailed by q (and therefore the alternatives that are not entailed by q are false). Notice that this notation is slightly different from Krifka's rule (28), which says there is no stronger (entailing) scalar alternative which is true, and Chierchia's (30), which states the stronger alternative is negated. However, as Chierchia points out, the different notations are equivalent.

Applying the new notation, the new definition of Strong Application (the counterpart of (33)) is (36):

(36) $[\![\alpha\ \beta]\!]_S = \{ \quad [\![\alpha]\!]_S([\![\beta]\!]_S) = [\![\alpha]\!]_S(O_C\ [\![\beta]\!])$ if $[\![\alpha]\!]$ is not DE. $\}$
$O_C\ [\![\alpha]\!]([\![\beta]\!])$ otherwise

This is equivalent to (33) above (combined with the old rule (30) and the Strength Condition (31)), with the conceptual advantage that the difference between non-DE and DE-environments is more straightforward here: the level (of embedding) at which O is applied. A nice result is that it is the same operator that accounts for the presence of embedded SIs in UE-contexts (where O is applied in embedded position), as for the absence of SIs with a top-item of a scale (where applying O is vacuous), as for the indirect SIs on a higher level in DE-contexts (where O is applied on the higher level, i.e. higher up in the bottom-up derivation, higher up in the syntactic tree).

Summing up, I have presented two accounts of SIs which are based on focus semantics. Krifka (1995) took the ordered <B,F,A> triples of the Structured Meaning approach and proposed a rule to derive SIs of scalar terms in focus. Chierchia (2004, 2006) adopted the view of Alternative Semantics that every expression has another value based on its alternatives, replacing the focus alternatives with scalar alternatives and proposing a rule (and later an exhaustivity operator) similar to Krifka's, in which stronger scalar alternatives are ruled out. Also, he showed how function application of alternative meanings, proposed for focus in Alternative Semantics, can be applied to scalar alternatives to predict the right SIs in different structural environments.

However, neither Krifka nor Chierchia makes predictions about contextual properties above the sentence-level. Although the account of Chierchia makes testable predictions about how structural, sentence-level properties (such as downward entailment) affect SIs, it does not make predictions based on the wider context, as Chierchia tries to account for SIs in what he calls *neutral* contexts:

> 'The claim is that there are situations in which (standard) implicatures are by default present and situations in which they are by default absent, and such situations are determined by structural factors. By default interpretation, I simply mean the one that most

> people would give in circumstances in which the context is unbiased one way or the other.' (Chierchia 2004, p. 51)

He acknowledges that if the context is not neutral, SIs that are derived by his system can be cancelled, and SIs that were not calculated can be accommodated. However, he does not specify which contextual properties are responsible for this (I return to Chierchia's account and this issue in section 5.2 of chapter 8). In the next section, I will discuss two proposals of SIs that do make predictions as to in which wider contexts SIs arise and in which they are typically absent. Even though these theories are not derived directly from the semantics of focus in the sense that the proposals in this section are, their predictions are based of the focus-background distinction.

4. Scalar implicature predictions based on the focus-background distinction

In this section I discuss two views that predict presence or absence of SIs based on the focus-background distinction. In 4.1, I present the account of Van Kuppevelt (1996), who relies on intuitions about questioning processes in discourse to derive the claim that SIs only arise in the focus part of the sentence. After that, I discuss the account of Van Rooij (2002), in which the focus dependency of SIs follows from the derivation of SIs by exhaustification of answers.[5] While both theories are set in different theoretical frameworks, they make the same strong claim that SIs will only arise if a scalar term is in the focus part of the sentence. This is the prediction that will be tested experimentally in this thesis. Finally, in section 4.3 I will explain why this prediction based on the focus-background distinction, which at first glance seems to be a sentence-level property, is actually reflecting a contextual property.

[5] Another theory which discusses the importance of focus for SIs is Hirschberg (1985). However, she describes the marking of focus (e.g. by syntactic or intonational means) as ways to express *salience*, which is responsible for triggering the SI. This is different from the current approach, in which focus is determined in terms of new information. Hirschberg also discusses old vs. new information, but actually makes an opposing prediction to the theories discussed here. According to her, givenness (so old information) might be associated with salience, which triggers SIs.

4.1 Van Kuppevelt (1996)

4.1.1 The framework

The account of Van Kuppevelt is set in the framework of *Discourse Topic Theory* (DTT) (Van Kuppevelt 1991). This is a model in which discourse structure is organized by an ongoing questioning process. The *topic* of a discourse unit (a sentence or a number of sentences) is determined by the explicit or implicit question it answers. This answer provides the *comment*. Van Kuppevelt says:[6]

> 'By definition, a topic T_p is *that which is being questioned* by means of a contextually induced explicit or implicit question Q_p. The corresponding comment C_p is provided by answer A_p. C_p is *that which is asked for* by Q_p.' (p. 396, his italics)

Semantically, the topic T_p is the intension of the topic term of the question, e.g. in (37Q), it is the intension of *(the one) who is laughing*.

(37) Q: Who is (the one who is) laughing?
A: Alan is laughing.

As the intension of a term is the set of possible extensions of this term, in a domain with only two people, Alan and Brian, the topic is (38), where *S* stands for *Situation*:[7]

(38) $T_1 = \{<S_1,\{Alan\}>, <S_2,\{Brian\}>, <S_3,\{Alan, Brian\}>\}$

The comment is the extension of the topic term in the actual situation, e.g. {Alan} for the answer in (37). If the answer uniquely determines the topic extension, T_p is closed off.

We immediately see the parallel between Van Kuppevelt's topic-comment distinction and Jackendoff's focus-background distinction. The comment of a sentence is that which is asked for by the explicit or implicit question, so it corresponds to the focus of the sentence, and the topic is the question predicate, so it corresponds to the background. To indicate this

---

[6] The subscript *p* is, as far as I can tell, meaningless.

[7] Notice that this is not a propositional account of questions and answers like Hamblin (1973), where a question is the set of all its possible (propositional) answers, but an *individualistic* one, where a question is the set of all possible term answers (Van Kuppevelt refers to Hausser 1983 i.a.).

correspondence, I will therefore use the compound terms *focus/comment* and *background/topic*.[8]

The questioning process that is assumed by Van Kuppevelt is the result of *indeterminacies* in the discourse (also called *question locations*). An indeterminacy / question location is 'a non-uniquely referring term which, because of its referential ambiguity, is made the subject of questioning and, as a consequence of this, becomes a topic expression' (p. 397). For the example in (37Q) at the moment of questioning the extension of the topic term is undetermined, the *topic range* (the set of possible extensional values) does not yet contain a unique value (see (38)). The reduction of the undeterminedness of the topic extension is realized by an answer to the question, like (37A). If a satisfactory answer is given (which means there is no epistemic limitation and the answer provides a unique determination of the topic extension), the topic range is reduced to one value, and as a result the undeterminateness of the question location disappears. As this undeterminateness was a necessary condition for topichood, the topic is closed off.

There are two functionally different types of topic-forming questions: *main / topic-constituting questions* and *subtopic-constituting subquestions*. The second type of questions are used in a process of subquestioning, which happens if an answer is unsatisfactory, so if it has not led to the reduction of the topic range to one value. This can be because the answer is *quantitatively* unsatisfactory (the comment value is incomplete), or *qualitatively* unsatisfactory (the comment value is not specific enough). With these notions in place, I will now turn to Van Kuppevelt's claims about SIs.

4.1.2 Focus-sensitivity of SIs on Van Kuppevelt's account

Van Kuppevelt's crucial claim is that SIs depend on whether the scalar term is in the comment/focus part of the sentence. He gives example (39).

(39) Q: Who has fourteen children?
A: Nigel$_{\text{Comment}}$ has fourteen children.

This example is based on the view that numerals are also SI-triggers, i.e. that the lexical meaning of fourteen is *at least fourteen*, and that because the numerals are ordered on an entailment scale <1,2,3,4...>, a numeral triggers

[8] Some authors have explicitly argued against collapsing *focus* with *comment* and *background* with *topic*, see e.g. Vallduvi (1990). The only reason I collapse these notions here is to clarify that for the simple examples I discuss in this work, Van Kuppevelt's theory can be translated into the notions *focus* and *background*.

the SI that the higher values on the scale (so 15, 16, 17, ... for *fourteen*) do not hold.[9] Therefore, the sentence *Nigel has fourteen children* should trigger the SI in (40), giving rise to the meaning in (41) for the sentence.

(40) Nigel does not have more than fourteen children.

(41) Nigel has exactly fourteen children.

Van Kuppevelt however claims the SI (40) does not arise in (39), as the scalar term *fourteen* is not part of the comment/focus. He writes:[10]

> '[...], it is not only doubtful but even highly unlikely that a scalar implicature is generated at all in these cases. If in (9) [here (39), AZ] a scalar implicature would have been induced as the result of the quantifying term *fourteen*, this would transform the semantically provided 'at least fourteen' interpretation of this term into 'exactly fourteen' implying that this term is no longer an indeterminacy and that, as a consequence, question induction is blocked. Example (9)' [here (42), AZ] illustrates that this prediction is wrong.

(42) Q1: Who has fourteen children?
A1: Nigel$_{\text{Comment}}$ has fourteen children.
<Q2>: <How many children does he have?>
A2: He has twenty$_{\text{Comment}}$.'

(p. 406)

Van Kuppevelt claims that the answer A1 in (42) gives rise to the implicit question Q2 asking for the exact number. This can only be the case if *fourteen* still gives rise to an indeterminacy, which is impossible if its interpretation is *exactly fourteen*, but possible if its interpretation is *at least fourteen*. He concludes that therefore the latter meaning, which is given by the semantics, has to be the right meaning here, and no SI is calculated for *fourteen* in A1. In A2 however, *twenty* is part of the comment/focus, as it is questioned by Q2. Therefore, the SI that John has no more than twenty children does arise in A2.

However, according to Van Kuppevelt A1 in (42) does give rise to another implicature: the term *Nigel* gives rise to the implicature that he is the

[9] The view of numerals as SI-triggers is highly debated these days (see e.g. Carston 1998, Breheny 2005), but this does not matter for the point being made here.
[10] Implicit questions are represented between < > by Van Kuppevelt.

only one (in the relevant domain) who has fourteen children. This is because *Nigel* has comment status here. Van Kuppevelt says:

> 'In other words, one of our criteria for implicature generation, and thus for scale activation, is that the inducing context must have comment function.' (p. 407)

So he claims that in order for a sentence with a scalar term to trigger an SI, the scalar term has to be in the comment/focus part of the sentence.

Van Kuppevelt mentions that other proposals in the literature have already stressed the importance of comment status for SIs. For instance Campbell (1981) said that cardinals only get an 'exactly' interpretation when they are in comment position. Fretheim (1992) also distinguished between scalar terms in focus and background, but makes different predictions. He claims that in the focus, the upper bound is an entailment, while in the background, the upper bound arises by an SI. Van Kuppevelt says the following about this:

> 'Apart from the fact that no evidence exists for the assumption that in the latter case an inference is actually generated, the preceding example (9)' [here: (42) AZ] illustrates that this possibility is ruled out by the simple fact that in such a case question induction is still an option.' (p. 407)

So Van Kuppevelt claims that for scalar terms that are not in comment position, the SI never arises, which can be witnessed from the fact that a following question about the exact value is still an option.

As Van Kuppevelt does not provide a theory about how SIs are derived (presumably these come about by Gricean reasoning), his claim about SIs only arising in comment/focus position crucially rests on the intuition about whether subquestioning is possible. In the next subsection, I discuss the theory of Van Rooij (2002), who makes the same prediction. However, on his proposal the prediction that SIs only arise in the part of the sentence that was questioned, follows from the way SIs are derived in this theory: by application of an exhaustivity operator to answers.

4.2 Van Rooij (2002)

4.2.1 Exhaustivity

Van Rooij (2002) adopts a proposal by Groenendijk & Stokhof (1984), who claim that in answers like in (43), a covert exhaustivity operator is applied to the answer, giving rise to the exhaustive interpretation that *only* John came to the party.

(43) Q: Who came to the party?
A: John.

Van Rooij argues this exhaustivity operator can be used to derive SIs. In honor of its inventors, he calls the operator exh^{GS}, the definition of which is given in (44):[11] [12]

(44) $\text{exh}^{GS} = \lambda T \lambda P \lambda w[T(P)(w) \wedge \neg\exists P'[T(P)(w) \wedge P'(w) \neq P(w) \wedge \forall x[P'(x)(w) \rightarrow P(x)(w)]]]$

In this formula, *T* is the denotation of the term answer (e.g. ⟦ John ⟧ in (43)) and *P* stands for the property underlying the wh-question (e.g. *coming to the party* in (43)). What exh^{GS} does, is pick out the *minimal elements* of the set of sets denoted by the term answer. For instance, assume that we are in a domain with three individuals (say John, Bill and Mary). Applied to the answer in (43), exh^{GS} then picks out the set of sets {{j}} from the GQ-meaning of *John* (which corresponds to {{j},{j,m},{j,b},{j,b,m}}. To see how his works, let's first simplify by giving an extensional version of exh^{GS}:

(45) $\text{exh}^{GS} = \lambda T \lambda P[T(P) \wedge \neg\exists P'[T(P') \wedge P' \neq P \wedge \forall x[P'(x) \rightarrow P(x)]]]$

If we apply this to the GQ-meaning of *John*, λPP(j), we get (46):

(46) $\lambda P[P(j) \wedge \neg\exists P'[P'(j) \wedge P' \neq P \wedge \forall x[P'(x) \rightarrow P(x)]]]$

---

[11] Notice that exh^{GS} is very similar to Krifka's rule and Chierchia's rule, given in resp. (28) and (34) above.

[12] Van Rooij proposes another exhaustivity operator later in the paper 'exh^{R}', which takes into account the difference between mention-all and mention-some questions. However, for the prediction of focus-sensitivity of SIs the basic operator $\text{exh}^{G\&S}$ suffices. In later work, Van Rooij also proposes more sophisticated exhaustivity operators (see Van Rooij & Schulz 2004, Schulz & Van Rooij 2006). One of these, the eps-operator of Van Rooij & Schulz (2004), will be discussed chapters 8 and 9 of this dissertation.

This formula denotes the set of sets that have *j* as a member and for which it holds that there is not a different set P' which has *j* as a member and is a subset of it. The set of sets which satisfies these requirements is the set of sets {{j}}. So for (43), the extension of the question predicate *coming to the party* is {{j}}, and we derive that only John came to the party, which matches our intuitions. Van Rooij demonstrates that in our domain with John, Bill and Mary, if the answer to the question in (43) would have been *John and Bill*, exh^{GS} correctly picks out {{j,b}}, and for *a man*, it correctly gives {{j},{b}}.

A good example of how exh^{GS} can account for SIs, follows from its application to a disjunctive term. A sentence containing *A or B* is usually considered to trigger the SI *it is not the case that A and B*, based on the entailment-scale <or, and> (see chapter 1). If we apply exh^{GS} to a term containing disjunction, we see that exh^{GS} gives us this result straightforwardly. For instance, if the answer to the question in (43) is *John or Bill*, we apply exh^{GS} to its meaning $\lambda P[P(j) \vee P(b)]$, which is the set of sets which contains at least one of John and Bill: {{j},{b},{j,b},{j,m},{b,m},{j,b,m}}. If we apply exh^{GS} to this meaning, we get (47):

(47) $\lambda P[(P(j) \vee P(b)] \wedge \neg \exists P'[(P'(j) \vee P'(b)) \wedge P' \neq P \wedge \forall x[P'(x) \rightarrow P(x)]]]$

(47) denotes the set of sets that have at least one of *j* and *b* as a member, and for which it holds that there is not a different set P' which has at least one of *j* or *b* as a member and which is a subset of it. The only sets in $\lambda P[P(j) \vee P(b)]$ which satisfy this condition are the set that contains only John, {j}, and the set that contains only Bill, {b}. So the set of sets that is picked out by exh^{GS} is {{j},{b}}. Crucially, the sets {j,b} and {j,b,m} are excluded by exh^{GS}, so we derive that it is not the case that John and Bill came to the party. This way the SI is derived by application of exh^{GS}.

So Van Rooij's proposal is that SIs are derived by a mechanism of exhaustification of answers by application of exh^{GS} to term answers. However, SIs are usually associated with whole sentences. Groenendijk & Stokhof (and Van Rooij) point out that for constituent questions like (43), the approach extends straightforwardly to sentential answers. If e.g. the answer to the question in (43) would have been the sentential answer *John came to the party*, and we apply exh^{GS} to this, we get the result in (48):

(48) $CAME(j) \wedge \neg \exists P'[P'(j) \wedge P' \neq \lambda x CAME(x) \wedge \forall x[P'(x) \rightarrow \lambda x CAME(x)]]$

This says that John came to the party and that there is no other set P' of which John is a member, and which is a subset of the set of party-goers. In

order to satisfy this requirement, the set of party-goers has to contain only John. Therefore, just like in the term answer case, *John* is exhaustified from its GQ meaning {{j}, {j,b}, ...} to the meaning {j}. So for sentential answers, exh^{GS} exhaustifies the constituent that could have been the term answer (*John*), but leaves the rest of the sentence, the question predicate *came to the party*, as it is (e.g. the denotation of *party* is not exhaustified).[13] This brings along an interesting prediction.

4.2.2 Focus-sensitivity of SIs on Van Rooij's account

Reconsider Jackendoff's focus-background distinction, where the focus was defined as the new information and the background as the old information. Obviously, the part of a sentential answer that could have been the term answer always corresponds to the focus, as it is the new information provided by the answer. Correspondingly, the rest of the sentence corresponds to the background, as it is the old information which was already given in the question. This generalization is known as Question-Answer Congruence for Focus: 'The position of focus in an answer correlates with the questioned position in wh-questions' (Rooth 1996, p. 271). So whether or not exh^{GS} will exhaustify a constituent of a declarative sentence, and therefore whether or not an SI will arise if a scalar item is present, depends on whether the constituent in which the scalar term appears, is part of the focus. This is the same prediction that was made by Van Kuppevelt above.

Van Rooij provides example (49), similar to Van Kuppevelt's (42) above.

(49) Q: Who has two children?
A: John has two children.

Just like Van Kuppevelt, Van Rooij claims the SI of the numeral is absent in (49), i.e. according to the answerer, John might as well have five children. Van Rooij's exhaustivity view can account for this straightforwardly: exh^{GS} picks the minimal set of sets of the focus *John*, leading to the exhaustivity inference that according to the speaker John is the only one (in the relevant domain) who has two children. However, exh^{GS} leaves the background *has*

[13] This is formulated somewhat sloppily. Actually, the extension of the question predicate (*came to the party*) is exhaustified, by reducing it to the minimal elements of the set of sets denoted by the term answer. However, I wanted to cash out the intuition that the action happens at the term answer, turning the meaning of *John* into *only John*.

two children, the part that contains the (alleged) scalar term, in tact. There is no picking out of a minimal set of sets going on there. Therefore, the sentence will not get the SI associated with *two*. By contrast, if the sentence would have been an answer to *How many children does John have?* the SI would have arisen on Van Rooij's view, because the scalar term *two* would have been the focus, and exh^{GS} would exhaustify it, giving rise to the SI.

The fact that the approach of Van Rooij crucially depends on the focus-background distinction is made more obvious in Van Rooij & Schulz (2004) (VR&S), where they provide this version of exh^{GS}:

(50) $exh(F,B) = F(B) \wedge \neg\exists B' \subseteq D: F(B') \wedge B' \subset B.$

VR&S say *exh* takes as arguments: '(i) the predicate *B* of the question, and (ii) the meaning of the term answer, or focus, *F* to the question.' (p. 498). So here term answer and focus are both captured by *F*, and (although this is not explicitly mentioned), the letter *B* was probably used because the question predicate of a wh-question is the background of the answer. Notice the similarity between this version of exh^{GS} and Krifka's rule in (28), repeated below, although the formulas seem to be mirror images because in VR&S's semantics the focus applies to the background: F(B), while in Krifka's this is the other way around: B(F).

(28) $Scal.Assert(<B,F,A>)(c) =$
$\{i \in c \mid i \in B(F) \wedge \neg\exists F' \in A[([c \cap B(F')] \subset [c \cap B(F)]) \wedge i \in B(F')]\}$

Although this is not made explicit in his paper, also Krifka's rule (28) predicts SIs only to arise if a scalar term is part of the focus.

4.3 Focus dependency as a contextual property

We established that both the account of Van Kuppevelt (1996) and the account of Van Rooij (2002) make predictions about SIs based on the focus-background distinction. However, the focus-background distinction is a property at the sentence level, and as was explained in the introduction, we would like to get a grip on contextual properties above the sentence level. The crux is that unlike other sentence level properties, such as downward entailment or modality, the focus-background distinction is determined by the preceding context. After all, it is a distinction between old and new information, and whether information is old or new can only be established by looking at the wider context. Furthermore, we have seen that an account based on the focus-background distinction makes different SI-predictions for

one and the same sentence depending on the context it appears in, and therefore it has to be an account of a contextual property. Van Rooij & Schulz (2004) put it this way:

> 'Another pleasing property of an exhaustivity analysis of implicatures is that it predicts that it depends on the context, or question-predicate, whether we observe these inferences. If, for instance, the scalar term occurs in the question-predicate P instead of the focus F of the answer, as for instance in example (ii), no implicatures are predicted.
>
> (ii) A: Do you have some apples?
> B: Yes, I have some apples.
>
> 'This may account (at least partially) for the often cited context-and relevance-dependence of implicatures' (Van Rooij & Schulz (2004), p. 499-500)

A consequence of the focus sensitivity of SIs, is that in order for a sentence to receive an SI, it has to be an answer to a question. Van Kuppevelt and Van Rooij both explicitly make this claim:

> '[...] on each discourse level the generation of a scalar inference is determined by the explicit or implicit (sub)topic-forming question [...]' (Van Kuppevelt (1996), p. 403)
>
> '[...] almost all typical quantity implicatures can be alternatively analyzed on the assumption that assertions are exhaustified answers to questions.' (Van Rooij (2002), p. 9)
>
> 'One might object to our approach, saying that scalar implicatures arise even when a scalar term is not used to answer a corresponding question. I believe this objection is ungrounded.' (Van Rooij (2002), p. 13)

However, in normal conversation or discourse, not every declarative sentence is preceded by an explicit question to which it is an answer. Therefore, both Van Kuppevelt and Van Rooij rely on the possibility that a question can also be implicit in the context. Van Rooij says:

> 'Thus, our exhaustification approach towards implicatures predicts that they depend on the topic being addressed. This topic can be an explicitly stated question or an implicit issue that is somehow relevant in the discourse.' (Van Rooij 2002 p. 13)

This corresponds to the view of Van Kuppevelt that discourse gives rise to implicit contextual questions which are answered as the discourse proceeds (see also Roberts 1996).[14] For instance, assume I just told you that last week after class, I told each of John, Bill and Mary that if they would hand in their homework early, they would get a higher grade. I can now utter (51):

(51) Mary handed in her homework early.

Even though in this case (51) is not an answer to an explicit question, it still gives rise to the exhaustivity inference that out of John, Bill and Mary, Mary is the only one who handed in her homework early. In this case, the question *Who handed in early?* was implicitly triggered by the preceding context. As a result, *Mary* is the focus of (51), and exh^{GS} applies to it, giving rise to the exhaustivity inference. Notice that exh^{GS} does not apply to constituents in the background, such as *her homework*. Even though it is not salient in this discourse, Mary could have just as well handed in something else besides her homework (say, her term paper). Also, the fact that *Mary* is the focus of (51) in this case, is reflected by the fact that this constituent receives the main stress of the sentence, which is not the default stress contour of the sentence. So by allowing for the question to be implicit, Van Kuppevelt's and Van Rooij's generalization that SIs only arise if a scalar term is in comment/focus position is extended to declarative sentences that are not direct answers to explicit questions.[15] Summing up, on Van Kuppevelt's and Van Rooij's accounts the presence or absence of SIs always depends on something bigger than just the sentence the scalar term appears in: it depends on the explicit or implicit question of the preceding context that the sentence is an answer to. In the next section I summarize the predictions of Van Kuppevelt and Van Rooij.

[14] Van Kuppevelt describes an implicit question as follows: 'implicit questions are defined as those questions the speaker anticipates the addressee asking as the result of the preceding context.'

[15] A term often used for the most salient question at a certain point of an ongoing discourse is *Question Under Discussion* (QUD), see Roberts (1996) for a formal implementation of this term. In earlier work (e.g. Zondervan, 2009), I used *QUD* to refer to both an explicit question and the most salient implicit question of a discourse. To avoid confusion over this, I will avoid this term in this work, and just refer to the two types of questions as *explicit* and *implicit*. I will also explicitly address the salience of the implicit question when this is relevant.

4.4 Experimental predictions of the focus-sensitive SI theories

Both the account of Van Kuppevelt (1996) and the account of Van Rooij (2002) are not directly derived from the semantics of focus, such as the accounts of Krifka (1995) and Chierchia (2004, 2006), but both do assign a key role to the focus-background distinction in predicting when SIs will arise. Therefore, these theories can be labeled *focus-sensitive* theories of SI. Both make the explicit claim that whether or not an SI arises depends on what part of the sentence the scalar term appears in, relative to the question it an answer to. This question can be either explicitly present in the context, or it can be implicit. If the scalar term is in the focus, the sentence will give rise to an SI, while if it is in the background, it will not.

The focus-sensitivity of SIs is a very testable prediction. One of its nice features is that for one and the same sentence, it predicts two different interpretations (one with SI and one without SI), depending on the context it appears in. Consider for instance (52) and (53), also given in chapter 1, in which the same sentence containing the scalar term *or* is an answer to two different (contextual) questions:

(52) Q: What did Harry bring?
A: Harry brought bread or chips$_F$.

(53) Q: Who brought bread or chips?
A: Harry$_F$ brought bread or chips.

Only in (52) is the SI (*it is not the case that Harry brought bread and chips*) predicted to arise, as there it is part of the focus. The same sentence should not get an SI in (53), as there the scalar term *or* is part of the background. In the next chapters, I will present a series of experiments that were set up to test these predictions.

In chapters 3 and 4 I present off-line experiments that tested these predictions. The experiments in chapter 3 employ the TVJT-paradigm, where question-answer pairs like (52) and (53) were explicitly given following a story in which both disjuncts were true (e.g. Harry brought both bread and chips), and participants were asked whether the answer that was given was true. I also tested whether the predictions were confirmed when the question was left implicit, and stress was used to manipulate the focus-structure of the target sentences. In chapter 4 I introduce a new paradigm to test SIs, the PWJT. In these experiments participants were explicitly asked about the possibility of the *A and B* situation, after stories containing question-answer pairs like (52) and (53). In the experiments in this chapter

another contextual property that possibly affects SIs was also tested: the contextual relevance of the stronger scalar alternative (*A and B*).

In chapter 5 and 6 I turn to on-line processing experiments. Chapter 5 discusses previous processing studies on SIs, which suggest that SIs bring along a processing cost. Based on this prediction I conducted two self-paced reading experiments in which reading times on segments containing *or* was compared between focus and background, using stories containing pairs like (52) and (53). These are presented in chapter 6. Finally, in chapter 7 I return to the TVJT, to test the predictions of focus-sensitivity of SIs on another scalar term in different conditions.

CHAPTER 3: ARE SIS FOCUS-SENSITIVE? THE TVJT EXPERIMENTS

1. Introduction

The goal of the series of experiments that are presented in this chapter was to test the predictions made by the theories of Van Kuppevelt (1996) and Van Rooij (2002) that Scalar Implicatures (SIs) are sensitive to focus. Both accounts predict that SIs only arise if a scalar term occurs in the part of the sentence that is the information focus. Following Jackendoff (1972), the information focus of a sentence is the part that contains the new information (versus the rest of the sentence, which contains the old information and is called the *background*). The focus structure of a sentence is determined by the explicit or implicit question in the context that the sentence is an answer to. Therefore, one and the same sentence will give rise to an SI if it is an answer to the one question, but not if it is an answer to another question. For example, the sentence *Harry brought bread or chips* can give rise to the SI *Harry did not bring bread and chips*, based on the entailment scale <or, and> (see chapter 1). According to the focus-sensitive accounts of SIs, this SI will only arise in (1), and not in (2):

(1) Q: What did Harry bring?
A: Harry brought <u>bread or chips</u>$_F$.

(2) Q: Who brought bread or chips?
A: <u>Harry</u>$_F$ brought bread or chips.

In (1) *or* is in the part of the sentence that was questioned (the direct object). Therefore, *or* is in the part that contains new information, the information focus, and the SI is predicted to arise. However, in (2) the subject is questioned. Since the scalar term *or* is not in this part of the sentence, it is not part of the focus and no SI is predicted.

In Experiments 1 and 2 I tested whether this prediction is borne out if the focus structure of the target sentence is determined by an explicit question preceding the target sentence, as in (1) and (2). In Experiment 1 I also manipulated the context to fit this question, while in Experiment 2 the only difference between the two conditions was the explicit question itself. In Experiment 3 I investigated whether the focus-sensitivity of SIs extends to situations in which the contextual question is implicit, but the focus structure of the target sentence is reflected by its stress pattern.

The experimental paradigm that was used in Experiments 1-3 was the *Truth Value Judgment Task* (TVJT).[1] This is the most widely used experimental task to access SIs, probably due to the reasonable simplicity of the task.[2] Therefore, I will start the investigation with this paradigm. The typical setup of a TVJT experiment on SIs is that participants are presented with a story. In the story a situation is described (or acted out), which could be described by the stronger scalar term, e.g. *A and B* is the case. Then a target sentence is presented (in acquisition studies often uttered by a puppet), and participants are asked to judge whether this sentence is true, or alternatively whether the puppet 'said it well'. However, the target sentence typically contains a weaker scalar term, e.g. *or*. Now, if the participant calculates the SI for the target sentence (she interpreted *A or B* as *A or B but not both*), the sentence becomes false relative to the story. However, if the participant does not calculate the SI and she takes *A or B* to mean *A or B and possibly both*, the target sentence is true. Therefore, the true/false judgment indicates whether or not the SI was calculated by the participant.

As is the case for all experiments presented in this dissertation, Experiments 1-3 were conducted in Dutch, with adult native speakers of Dutch as participants. However, the data should be comparable to data obtained for other languages such as English, as pragmatic inferences like SIs are based on cooperative behavior of language users, which I assume is not language-specific.[3] Also, the hypothesis that I will test is based on the distinction between old and new information, which is also hardly language-specific. Throughout the thesis I will only provide English counterparts of the experimental items. The interested reader is referred to the Appendix for the original Dutch items (and English translations).

---

[1] I use the acronym *TVJT* throughout this thesis to refer to the paradigm in which I ask adult speakers for truth value judgments of an *A or B* sentence after an *A and B* story. However, it has been pointed out to me that this acronym is used mostly for the version of the task described by Crain & Thornton (1998) used in acquisition, which involves an act-out task and a puppet etc. I hope this will not lead to confusion.

[2] Another reason could be that SIs are a popular topic in language acquisition research, as they seem to be notoriously hard for children to handle (see e.g. Chierchia et al. 2001, Papafragou & Musolino 2003). As the TVJT is suitable for young children this could explain its popularity among researchers investigating SIs.

[3] Of course, subtle lexical differences of certain scalar terms between languages could complicate comparison, e.g. English *some* has at least two counterparts in French (*quelques* and *certains*) and Dutch (*enkele* and *sommige*), which do not necessarily behave alike in triggering SIs (see section 1 of chapter 7 for discussion). Therefore, comparison of absolute numbers of studies conducted in different languages is risky.

The layout of the chapter is as follows: In sections 2-4 I present Experiments 1-3. Each section will be closed off with a short discussion of that particular experiment. Sections 5-7 are discussion sections which relate to all three experiments. In section 5 I will consider three issues that might have decreased the observed effect. In section 6 I critically consider the suitability of the TVJT to test the current hypothesis and to assess SIs in general. Finally, in section 7 I turn to two alternative hypotheses from the psycholinguistics literature about how focus might affect SIs and I consider their relation to the current experiments. Section 8 provides a summary of the discussion and conclusions.

2. Experiment 1: Focus through explicit questions and contextual support

2.1 Setup and items

In Experiment 1 I used a written version of the TVJT. The experiment was set up to test the focus-sensitivity of SIs where the focus structure of the target sentence is determined by an explicit question. Therefore, contrary to a standard TVJT experiment, the story was not followed by a target sentence in isolation but by a dialogue between two speakers. Speaker A asked a question about the story and Speaker B answered. Participants were asked to judge whether Speaker B's answer was true or false. The focus structure of the target sentence was manipulated by varying Speaker A's question between conditions, like in (1) and (2) above. Example items of the two conditions are given in (3) and (4):

(3) Condition 1: focus
Katja was searching for marine animals on the beach near her grandparents' house. She had promised her grandfather to find some beautiful animals. He had said that if she would find an oyster, she would get ten bucks. Katja soon found a crab. Not much later she also found a starfish. But no matter how hard she looked, she didn't find an oyster.

A: "What did Katja find?"
B: "Katja found a crab or a starfish."

Is the answer of speaker B true or false? true / false

(4) Condition 2: non-focus
Katja and Birgit were searching for marine animals on the beach near their grandparents' house. Their grandfather had encouraged them both to go look for a crab or a starfish. He had promised them that the one who would find a crab or a starfish, would get ten bucks. After some searching Katja found a crab. Not much later she also found a starfish. Birgit couldn't find anything and had to return to the house empty handed.

A: "Who found a crab or a starfish?"
B: "Katja found a crab or a starfish."

Is the answer of speaker B true or false? true / false

In the focus condition in (3), *a crab or a starfish* is the focus of the target sentence, due to A's question about the direct object. Therefore, it is predicted that participants will calculate the SI and judge B's answer false in this condition. In the non-focus condition in (4) however, A's question is about the subject, so *or* is not in the focus part of the sentence. Hence, no SI is predicted and participants are expected to judge B's answer true.

As illustrated by (3) and (4), I designed the contexts so that they would fit well with the question asked by speaker A. The goal of these manipulations was to make sure the question made sense in relation to the story, as a mismatch between question and story might lead participants to disregard the question and judge the sentence on its own, possibly assigning it a different focus structure than the one that was triggered by the question.

The first contextual manipulation was using different set sizes in the two conditions. The rationale behind this was the following: if a story contains one person, and a lot of objects she can choose from, the question that arises naturally is the question about the set of objects: *What did she choose?* On the other hand, if a story contains a lot of people and only one object, the question that arises naturally is about the set of people: *Who got/took the object?* So a bigger set versus a set of one triggers a question about the members of the bigger set. After all, a question about the set of one never arises, as it has only one member. In the focus condition, exemplified in (3), there was always only one person and three objects. Therefore, the natural question that arises is about what that person found, which is speaker A's eventual question. In this condition, the *who*-question is not interesting as only one character was introduced. Contrastively, in the non-focus condition, exemplified in (4), I introduced two people, allowing for the question about the subject. Unfortunately, due to the fact that two objects are needed to form a disjunction, it was impossible to have only one object.

Another effect of the use of different set sizes was that it created a *contrast*. In the focus condition, two objects were found (a crab and a starfish), and one was not (an oyster), while in the non-focus condition one person found a crab or a starfish (Katja), and one didn't (Birgit). These contrasts made a question about the object more interesting in the focus condition, and a question about the subject in the non-focus condition. Therefore, these contrasts also supported the question that speaker A asked.

The final contextual manipulation was that in the third sentence of both conditions a conditional was introduced to support speaker A's question by providing a reason why this question was relevant: In (3) the question *What did Katja find?* is relevant as she was promised a reward of ten bucks for finding an oyster. Similarly, in (4) the question *Who found a crab or a starfish?* is relevant, as the person(s) who did gets ten bucks.

I deliberately avoided using the stronger alternative with *and* in the stories (e.g. *After some searching Katja found a crab and a starfish*). In all items I divided the description of the two objects being 'verbed' (in this case: found) over two sentences, connected by something like *also*. I wanted to make sure rejections would be based on a comparison of the interpretation of the sentence to the situation, not on a surface discrepancy between the target sentence and a sentence in the story.

2.2 Design

6 story pairs like the one in (3)-(4) were created, and 14 fillers. The test items are given in Appendix 1. Four experimental lists were created as follows: List 1a was created by picking three items in condition 1 and three items in condition 2 (from different story pairs). List 1b differed from list 1a on condition: every item in list 1a was replaced by the same story in the other condition in list 1b. List 2a differed from list 1a on order (both items and fillers). It contained the same items in the same conditions as list 1a but in another order. List 2b differed from list 2a on condition, not on order. As a result, list 2b differed from list 1b on order and not on condition. The first test item of List 1a and list 2a was a non-focus item, while it was a focus item in list 1b and list 2b. The lists were distributed evenly over participants. Each list contained three items per condition, interspersed with 14 fillers, making a total of 20 stories per list. There were always at least two fillers between two test items. The 14 fillers were comparable stories with comparable dialogues, in which the given answer was true or false irrespective of SI. Some fillers contained *or* in the target sentence, and some did not. Half of the fillers contained a question about the subject and half

about the direct object. The contrast sets in the stories matched these questions. Half of the fillers was true and the other half was false.

2.3 Participants and procedure

The experiment was a web-based questionnaire built with the experimental software WWSTIM (Veenker 2000). 37 people were recruited via email and filled out the questionnaire on their own computers. All were adult native speakers of Dutch with no prior knowledge of the topic. Most of them were students or had a university degree (in a non-related field). Participants were instructed they would read a story followed by a dialogue between Speaker A and Speaker B, where Speaker A would ask a question about the story and Speaker B would answer. Their task was to judge whether Speaker B's answer was true or false. The stories were presented one by one on the screen, including the dialogue, the question *Is the answer of Speaker B true or false?* and the answering options 'true' and 'false'. These could be selected by clicking the right option. It was impossible to select both options, but correcting an answer was possible. Pressing the 'next' button allowed the participants to proceed to the next item. Participants were instructed to judge whether the sentence was true or false, and not whether it was weird or ugly. Most participants completed the experiment within 15 minutes, with an average of 10 minutes.

2.4 Results

All participants scored over 75% correct on the fillers so no participants were excluded from the analysis. One participant was excluded due to a number of double values, as a result of repeatedly pressing the 'back' button in the web browser. The results of the remaining 36 participants were included in the analysis. The results are given in Table 1:

Table 1. SI-rates Experiment 1

Condition 1 (focus)	Condition 2 (non-focus)
73%	55%

The percentage of 'false' answers in the focus condition was 73%, versus 55% in the non-focus condition. This difference was significant over subjects: Wilcoxon signed rank test gives $z = -3.26$, $p < 0.001$ (one-tailed), with effect size $r = -0.38$, and over items: $z = -2.21$, $p = 0.014$ (one-tailed), r

= -0.64.[4] Therefore, the results support the focus-sensitivity of SIs: more SIs were calculated when *or* was in the focus part of the sentence, than when it was part of the background.

The items analysis showed that for all stories more SIs were calculated in the focus condition (see Appendix 1 for the SI-rates per item). An interesting pattern arises from the subjects analysis when we consider the distribution of participants over the four possible types of behavior on the conditions. This distribution is given in Table 2, where 'more SIs' means at least one 'false' answer more, 'SIs in both conditions' means 3/3 'false' answers in both conditions, and 'SIs in neither' means 0/3 'false' answers in both conditions.[5]

Table 2. Distribution of participants Experiment 1

more SIs in foc.	SIs in both cond	SIs in neither	more SIs in nonf
42% (15)	33% (12)	17% (6)	8% (3)

We see that almost half of the participants (42%) distinguished between the two conditions in the predicted direction: they calculated more SIs in the focus condition. Crucially, only 8% showed the opposite pattern. In other words, 83% of the participants that distinguished between the two conditions did so in the predicted direction. However, a large group (50%), did not distinguish between the conditions and either calculated SIs in both conditions, or in neither condition.

2.5 Discussion

The results support the hypothesis that SIs are sensitive to focus, and more specifically that more SIs are calculated if the scalar term is part of the focus of the sentence. However, the difference is not as big as predicted by the

[4] The SI-rates of participants was not normally distributed in either condition (Kolmogorov-Smirnov test and Shapiro-Wilk tests both $p<0.001$ for both conditions, as well for the difference between the conditions), so we have to resort to non-parametric tests.

[5] In Zondervan (2007), I report the following distribution: Only SIs on focus: 31% (11), SIs on both: 47% (17), SIs on neither: 19% (7), only SIs on non-focus: 3% (1). That was based on the following cut-off points: 0 or 1 'false' answers out of 3 items of a condition was considered 'no SIs', while 2 or 3 'false' answers out of 3 was considered 'SIs'. So for instance 0 'false' answers out of 3 on the focus condition and 1 'false' answers out of 3 on the non-focus condition is grouped under 'more SIs in non-foc' here, while in Zondervan (2007) this participant was grouped under 'SIs on neither'.

theories of Van Rooij and Van Kuppevelt. These accounts predict that there should be close to 100% SIs in the focus case, and the SI-rate in the non-focus case should be close to 0%.[6] Instead, we observed 73% vs. 55%. In chapter 1 I already mentioned that previous experimental work has shown that SIs are hardly as well-behaved as predicted by the theory. Here I will briefly consider one experiment in more detail, which illustrates that also for structural factors that affect SIs the experimental data are far from black and white.

A study by Chierchia et al. (in press) compared the interpretation of *or* in two non-downward entailing (non-DE) contexts (a simple assertion and the consequent of a conditional) to two DE-contexts (the antecedent of a conditional and the restrictor of *every*). Examples are given in (5)-(8):

(5) Non-DE context 1 (simple assertion):
Jeremy is a child or foreign. He must fill out a form.

(6) Non-DE context 2 (consequent of conditional):
If someone must fill out a form, he is a child or foreign.

(7) DE context 1(antecedent of conditional):
If Jeremy is a child or foreign, he must fill out a form.

(8) DE context 2(restrictor of *every*)
Everyone who is a child or foreign must fill out a form.

As in a DE-context the strength of the scalar terms is reversed (e.g. *not A or B* rules out more possible situations than *not A and B*), *or* is stronger than *and* in a DE-environment, so it should not trigger the *not and*-SI (see chapter 1). Therefore, close to 0% SIs are predicted for (7) and (8). However, Chierchia et al. found 59% SIs for (7) and 42% SIs for (8). The results of the non-DE contexts (5) and (6) were also not close to 100%, but respectively 67% and 68%. So even for a clear structural property such as downward monotonicity, of which it is generally agreed by theoreticians that it cancels (or flips) SIs, the absolute numbers observed in an experiment are much less clear. This shows that with pragmatic inferences such as SIs we cannot rely too much on the absolute numbers, but we have to look at the differences

[6] Both Van Kuppevelt and Van Rooij indicate other sources of SI-absence. Van Kuppevelt mentions that SIs do not arise if the context triggers a process called 'topic weakening', and Van Rooij points to so called 'mention-some' questions as a cause for SI-absence. However, as these special circumstances are not present in the experimental items, an SI-rate close to 100% is predicted in the focus condition.

between conditions. In section 5 I will discuss some other explanations for why the difference is not as big as predicted.

As I said above, the contexts leading up to the question of speaker A were manipulated to make the question of speaker A fit the context. A downside of varying this material between conditions, is that it might have introduced unwanted differences between the conditions. For instance, one might point to the fact that in the non-focus condition, there is an earlier occurrence of *or*, namely in the third sentence: *He had promised them that the one who would find a crab or a starfish, would get ten bucks* in (4). In this sentence *or* appears in a DE-environment (the restrictor of *the one*) and we know that DE-environments cause SIs to be blocked or reversed. It is possible that participants somehow held on to this meaning of *or* when they interpreted the target sentence. This would be an alternative explanation for why fewer SIs were observed in this condition than in the focus condition, in which there was no earlier mention of *or*. To control for interfering factors like this one I decided to conduct a follow-up experiment in which only the explicit question was varied between conditions and the rest of the story was kept constant. This way we can make sure that no other differences between the stories are responsible for an effect. This experiment is presented in the next section.

Another possible interfering factor in Experiment 1 which I wanted to control for in Experiment 2, was the status of the dialogue. It was unclear to the participants who speaker A and speaker B were, and what their relation to the story was. As a result, the epistemic states of the speakers might have been unclear, which might have made the task of judging whether Speaker B's answer was true unnecessarily complicated. Therefore, I decided to make the speakers and the dialogue part of the story in the follow-up experiment, so their identity and their epistemic state was clear from the story. For this I switched to a dialogue in indirect speech.

3. Experiment 2: Focus through the explicit question only

3.1 Setup and items

Experiment 2 was set up to isolate the effect of focus triggered by an explicit question on SIs. The stories of the two conditions were kept constant up to the explicit question, so nothing else in the story could be held responsible for an effect. To allow for comparison to Experiment 1 again the TVJT was used. Example items of the two conditions are given in (9) and (10), where differences between the conditions are marked in boldface for easy reference. Naturally, there was no such marking in the experimental items.

(9) Condition 1: focus
Julie and Karin were searching for marine animals on the beach. After some searching Julie found a crab. Not much later she also found a starfish. Unfortunately, Karin didn't find anything.
When **Karin** returned, her mother asked **what kind of marine animals Julie had found**.
Karin answered that Julie had found a crab or a starfish.

Is Karin's answer true? true / false

(10) Condition 2: non-focus
Julie and Karin were searching for marine animals on the beach. After some searching Julie found a crab. Not much later she also found a starfish. Unfortunately, Karin didn't find anything.
When **they** returned, their mother asked **who had found a crab or a starfish**.
Karin answered that Julie had found a crab or a starfish.

Is Karin's answer true? true / false

As the stories up to the explicit question were kept constant over conditions, stories had to be used that allowed for an explicit question about the subject, as well as an explicit question about the direct object. Therefore, it was necessary to introduce two characters in both conditions, otherwise the *who*-question in the non-focus condition would not make sense. After introducing the two characters and what they were doing, the story simply provided the information needed for the TVJT, e.g. that Karin found both a crab and a starfish, and Julie found neither.

To make the dialogue and the speakers part of the story, I used indirect speech. The explicit question was now asked by a character in the story (the mother in (9) and (10)), of whom it could be assumed she was not present during the search. The answer was given by one of the characters who was present at the search. Therefore, contrary to Experiment 1, there could be no confusion about the epistemic states of the speakers here: the person asking the question did not know what happened, and the person answering did. As a result of the use of indirect speech, the target sentence (*Julie had found a crab or a starfish*) was embedded in the matrix clause (*Karin answered*...). This does not make any difference for the focus structure of the target sentence though, so it does not affect our predictions.

The only thing that was varied between the two conditions was the explicit question that the target sentence was an answer to.[7] Therefore, the problem that a difference in SI-rates between the two conditions might be attributed to something else than the focus difference, such as earlier use of *or* in a DE-environment, is avoided. If a difference between the two conditions is observed, it can only be due to the manipulation of the explicit question.

3.2 Design

Two story-pairs of Experiment 1 were adapted to the template of Experiment 2, and four new story-pairs were created, making a total of 6 story-pairs (see Appendix 2). This was necessary because not every item of Experiment 1 could be adapted to the template of Experiment 2 and still form a coherent discourse. The design of Experiment 2 was identical to that of Experiment 1. I created 4 lists from 6 story pairs in the same way as in Experiment 1. 14 new fillers were created similar to the test items, containing different characters and objects than the fillers in Experiment 1. Similar to Experiment 1 the fillers either contained *or* and were true or false irrespective of SI, or they did not contain *or*. Again, the numbers of fillers with a subject-question and with an object-question were matched, as well as the number of true and false fillers.

---

[7] Another difference between the two conditions was *When Karin returned* vs. *When they returned*. Consider what would have happened if *Karin* would be replaced by *they* in (1), or *they* by *Karin* in (2):

(1') When they returned, their mother asked what kind of marine animals Julie had found.
(2') When Karin returned, her mother asked who had found a crab or a starfish.

In (1') it seems the mother is addressing Julie, asking her: "What kind of marine animals did you find?" It would then be strange that Karin would answer this question. In (2'), the question of the mother seems odd, as the domain of *who* is unclear. Therefore, this difference between the two conditions was unavoidable. However, I do not think this difference is of any importance for our predictions. If it introduces any bias, it goes against our predictions: Perhaps in the focus condition Karin returned home before Julie stopped searching, and therefore she might not be totally informed about what Julie found. That would only lower the SI-rate for the focus condition, which goes against our predictions.

3.3 Participants and procedure

46 participants were recruited to participate in the experiment. Contrary to Experiment 1, the experiment was conduced on computers in the lab. The software that was used was however the same as in Experiment 1 (WWSTIM, Veenker 2000). All participants were undergraduate students of Utrecht University, and were paid for their participation. None of them had any prior knowledge of the topic. The procedure was the same as in Experiment 1, except that next to being provided with instructions on the screen, participants were also instructed by the experimenter. The instructions themselves were the same as in Experiment 1, except that participants were additionally reminded that there were no 'right' or 'wrong' answers, so they should follow their intuition. All participants completed the experiment within 20 minutes.

3.4 Results

None of the participants scored below 75% correct on the fillers so all were included in the analysis. The results of Experiment 2 are given in Table 3:

Table 3. SI-rates Experiment 2

Condition 1 (focus)	Condition 2 (non-focus)
67%	41%

Similar to Experiment 1, the SI-rate the focus condition (67%) was higher than in the non-focus condition (41%). Both rates are a bit lower than in Experiment 1, but the difference is roughly the same size. This difference was significant over subjects: Wilcoxon signed rank test gives $z = -4.01$, $p < 0.001$, with effect size $r = -0.42$, and over items: $z = -2.21$, $p = 0.014$, $r = -0.64$.[8] So again, the focus-sensitivity of SIs was observed.

The distribution of participants was also similar to that of Experiment 1. Consider Table 4, which is based on the same criteria as Table 2 above.

[8] The fact that the result over items is the same as in Exp 1 is no error, but a result of the ranking of the data in the Wilcoxon signed-rank test. Apparently, this ranking was exactly the same for the SI-rates of the items of Experiment 1 and Experiment 2.

Table 4. Distribution of participants Experiment 2

more SIs in foc.	SIs in both cond	SIs in neither	more SIs in nonf
50% (23)	17% (8)	26% (12)	7% (3)

Exactly half of the participants made a distinction between the focus condition and the non-focus condition in the predicted direction, while again almost none showed the opposite pattern (89% of the participants that distinguished between the conditions, did so in the predicted direction). Again, another large group (43%) did not differentiate between the two conditions. The items analysis showed that for all stories more SIs were calculated in the focus condition (see Appendix 2 for the SI-rates per item).

3.5 Discussion

The results of Experiment 2 confirm the focus-sensitivity of SIs. Because the two conditions only varied in the explicit question that was asked, we can conclude the effect has to be due to this, and cannot be caused by some external factor. Therefore, the results are a strong argument in favor of the focus-sensitivity of SIs.

The SI-rates were similar to those of Experiment 1, and although the difference was a bit bigger than in Experiment 1 (26% versus 18%), it was still not as big as predicted by the theory. Above, I already indicated that this is hardly ever the case with SIs, and in section 5 I will present some other possible explanations for this. However, there is one possible reason why the difference might be smaller than expected in Experiments 1 and 2 which I would like to address here, and that is the possibility of carry-over between conditions due to the within-subjects design.

The test items of both conditions were very similar. Both involved the use of *A or B* where *A and B* was the case. In this sense, the items stuck out, as none of the fillers contained this discrepancy (as I wanted to avoid the critical inference in the fillers). Participants may have recognized the similarity between the items, and as participants typically try to be consistent throughout the experiment, they might for that reason have judged the items of the two conditions the same. This would explain the reasonably large group of participants who did not distinguish between the conditions in Experiments 1 and 2. They deliberately did not distinguish between them, as they thought the items were of the same type.

Next to the fact that the items stuck out, results in the literature suggest that participants generally stick to their interpretation of a sentence with *or* throughout the experiment. In a reasoning study Noveck, Chierchia, Chevaux, Guelminger & Sylvestre (2002) presented participants with

questions containing *or* (*Is there a Q or an R?*), in which SIs are considered not to arise, as well as declarative sentences containing (*There is a Q or an R*), which can trigger SIs. 27 out of 32 participants did not change their interpretation of *or* between the two conditions. Moreover, of the participants that were presented with the question condition first, 31% calculated the SI for the question condition, while of the participants that were presented with the question condition after the declarative condition, 81% calculated the SI. Similarly, declarative sentences returned 87% SIs if they were presented first, but only 56% after the question condition. Similarly, a study on or by Chavallier, Noveck, Nazir, Bott, Lanzetti & Sperber (2008) showed a 50% difference between two conditions in a between-subjects experiment, and only 20% in a within-subjects replication, also pointing to the effect of carry-over. I return to this experiment in 7.2 below.

To see if the explanation of carry-over matches the results, consider Tables 5 and 6 below, where three rates are given for both conditions. The first row is the total SI-rate in that condition. The second row is the percentage of SI-answers of the first item of a condition participants encountered, irrespective of whether they had previously encountered an item of the other condition. The third row is the rate of SI-answers on the first item in a condition, where only the items were included that were not preceded by an item of the other condition (and hence were the first test item a participant encountered).

Table 5: Percentage SIs on all items and first items per condition Exp. 1

	focus	non-focus
all items	73%	55%
first item of that condition	75%	50%
only very first test item	71%	37%

Table 6: Percentage SIs on all items and first items per condition Exp. 2

	focus	non-focus
all items	67%	41%
first item of that condition	63%	35%
only very first test item	61%	26%

We see that in the focus condition, the percentage on the first items (second and third rows) are reasonably representative of the total percentage (first row). However, in the non-focus condition we see the total percentage is quite a bit higher than the percentage on the first items, especially if these were not preceded by a focus item. This suggests that there was carry-over

from the focus condition to the non-focus condition. It seems that once participants had encountered a focus item, they were more likely to judge a non-focus item false. The differences between the percentages in the final row of both tables provide a much bigger contrast between the two conditions for both experiments (34% and 35%).[9] In that sense, these numbers are more in line with the predictions of the theory, although they are still far from a 0%-100% contrast. To control for the effect of carry-over, I decided to use a between-subjects design in the next experiment.

The alternative explanation for the data of Experiment 1 that the interpretation of *or* in the non-focus condition was adopted from the *or* in a DE-environment in the story, is rejected by Experiment 2. Also when there was no previous mention of *or* in a DE-environment, the SI-rate for the non-focus condition was significantly lower than for the focus condition. However, there was another mention of *or* in the non-focus condition that was not present in the focus condition, the *or* in the explicit question. This was unavoidable as *or* had to be part of the old information in the target sentence. However, it is assumed that normally SIs do not arise in questions (see e.g. Noveck et al. 2002). Therefore, it might be argued that participants take the meaning of the VP from the question, which contains inclusive-*or*, and simply copy this meaning when they interpret the target sentence. As a result, fewer SIs arise in the non-focus condition than in the focus condition, in which the *or* in the target sentence was not repeated from the question. In Experiment 3, which I present in the next section, I left out the explicit question to control for this bias. In order to still be able to manipulate the focus structure of the sentence I used spoken stimuli and manipulated the sentence stress.

4. Experiment 3: Focus through implicit questions and stress

4.1 Setup and items

The goal of Experiment 3 was to replicate the focus effect without an explicit question, and in a between-subjects design. In Experiment 1 the explicit question was supported by contextual means, so a possibility would be to simply remove the explicit question from those items. After all, Van Kuppevelt and Van Rooij both claim the question that determines the focus structure of the target sentence does not necessarily have to be explicit, it can

[9] These differences are significant in a between-subjects analysis: Exp. 1: Mann-Whitney U = 98.50, p = 0.023 (one-tailed), Exp. 2: U = 172.50, p = 0.009 (one tailed).

also be triggered implicitly by the context. However, simply taking out the explicit question from the items of Experiment 1 was not an option. Consider the versions of the example items of Experiment 1 without an explicit question:

(3') Condition 1: impossible focus condition
Katja was searching for marine animals on the beach near her grandparents' house. She had promised her grandfather to find some beautiful animals. He had said that if she would find an oyster, she would get ten bucks. Katja soon found a crab. Not much later she also found a starfish. But no matter how hard she looked, she didn't find an oyster.

target: Katja found a crab or a starfish. true / false

(4') Condition 2: impossible non-focus condition
Katja and Birgit were searching for marine animals on the beach near their grandparents' house. Their grandfather had encouraged them both to go look for a crab or a starfish. He had promised them that the one who would find a crab or a starfish, would get ten bucks. After some searching Katja found a crab. Not much later she also found a starfish. Birgit couldn't find anything and had to return to the house empty handed.

target: Katja found a crab or a starfish. true / false

The problem is that due to the TVJT, in which a target sentence is checked with a story, the target sentence is not part of the discourse, while the predictions of Van Rooij and Van Kuppevelt about implicit questions are based on a sentence being part of a discourse. As a result, it is not clear in (3') and (4') that the target sentence should be considered to be an answer to a contextually triggered question. Furthermore, because the target sentence is not part of the discourse, there is no guarantee that participants will assign the right focus structure to it. We cannot make the target sentence part of the discourse as in Experiment 2, as that would entail introducing an explicit question.

Fortunately, there is a way to indicate the target sentence should be considered as an answer to a question and to unambiguously determine its focus structure, and that is by manipulating the main stress of the sentence. Remember that another reflection of the focus feature was that the main stress of the sentence falls on (a syllable of) the focus (see chapter 2). Therefore, if a constituent contains the main stress of the sentence, we know

it is (part of the) focus, and we know which question(s) it is an answer to. Therefore, in Experiment 3 I used spoken stimuli in which the main stress of the target sentence varied between conditions. In the non-focus condition, the subject received the main stress, while in the focus condition the direct object *A or B* received the main stress. Examples of target sentences are given in (11) and (12), in which capitals indicate the main stress.

(11) target sentence focus condition:
Paola took AN APPLE OR A PEAR from the fruit section.

(12) target sentence non-focus condition:
PAOLA took an apple or a pear from the fruit section.

The target sentences that were recorded were spoken by a phonetician. In the focus condition exemplified in (11), the whole NP *an apple or a pear* was accented by two H* accents (one per disjunct), while the rest of the sentence had low pitch. This indicated it was an answer to the question *What did Paola take from the fruit section?* Crucially, the phrasal tone of the disjunction was low (pitch went down between the two H* accents), unlike in alternative questions (such as *Do you want apple juice or orange juice?* see Pierrehumbert & Hirschberg 1990). Also, there was no pause after the first disjunct and no slowing down during the disjunction. The rest of the sentence was deaccented and had low pitch, which indicated it was the background. In the non-focus condition (12) the subject was accented with one H* accent, and the rest of the sentence had low pitch, indicating it was an answer to the question *Who took an apple or a pear at the fruit section?*

Notice that due to the PP that followed the direct object, the main stress in both conditions was different from the neutral stress pattern of the sentence, in which the main stress of a sentence typically falls on the rightmost (or most deeply embedded) constituent (Chomsky & Halle 1968, Cinque 1993). So by including a deaccented PP at the end of the sentence, it was ruled out that the target sentence had a wider focus. It could for instance not have been an answer to another question like *What did Paola do?*

Similar to Experiment 1 I also manipulated the context to support the question that the target sentence was an answer to. I again used different set sizes to make the question about the object more natural in the focus condition, and the question about the subject in the non-focus condition. In addition I changed the presentation of the situation. Consider the example items (13) and (14):

(13) Condition 1: focus
Paola was getting lunch at the cafeteria. At the fruit section there were bananas, oranges, apples and pears.
Bananas Paola didn't like, so she didn't take any.
The oranges looked a bit old, so she also didn't take any of those.
The apples looked nice and juicy, so Paola took one.
A pear Paola hadn't eaten in years, so for a change she decided to also take a pear.

"Paola took AN APPLE OR A PEAR from the fruit section."

(14) Condition 2: non-focus
Paola, Linda, Betty and Ginger were getting lunch at the cafeteria. At the fruit section there were only apples and pears.
Paola was a real health freak, so she took an apple. After some consideration she also took a pear.
Linda already got a lot of other food, so she decided not to take any fruit.
Betty didn't like fruit at all, so she ignored the apples and the pears.
Ginger considered taking a pear, but she wasn't sure whether she brought enough money, so she didn't take any.

"PAOLA took an apple or a pear from the fruit section."

In the focus condition in (13) four types of fruit were introduced, and only one character (Paola). This made the question salient what types of fruit Paola selected. The situation was given by considering the types of fruit one by one, saying of each of them whether it was selected by Paola. The types of fruit were always mentioned sentence-initially, either as a subject (*The oranges looked a bit old*), or by a topicalization structure which is very common in Dutch (*Bananas Paola didn't like*). The big set of fruits, the fact that they were considered one by one and the contrast between the fruits that were selected and the ones that were not, triggered the contextual question *What/which fruits did Paola take?* The focus structure of the target sentence matched this question, as the main stress on the direct object indicated this part of the sentence is the information focus.

Contrastively, in the non-focus condition in (14) four characters were introduced and only two types of objects (one was impossible as two disjuncts were needed). The story considered the characters one by one, saying of each of them whether or not they took an apple or a pear. The big contrast set of girls, the fact that the story considered them one by one and

the contrast between the girl that did take an apple or a pear and the ones that did not, triggered the contextual question *Who took an apple or a pear?* The focus structure of the target sentence, which was indicated by the main stress on the subject, matched this question.

4.2 Design

As we saw in Tables 5 and 6 above, the fact that the observed difference in Experiments 1 and 2 was smaller than predicted, might have been due to carry-over between the conditions. Therefore, Experiment 3 was set up as a between-subjects design. Another reason for this design was that it might become too confusing for participants to be presented with target sentences with different stress patterns in one experiment. Three story pairs like the one in (13)-(14) were created, see Appendix 3. The non-focus items were included in the non-focus version of the experiment, and the focus items in the focus-version. Both versions therefore contained three test items, interspersed with 7 fillers. In order to make the fillers similar to the test items, the non-focus version contained versions of the fillers which matched the 'many characters – few objects' template of the non-focus condition, and the focus version contained versions of the fillers matching the 'one character – many objects' template of the focus condition. In accordance with this all fillers in the non-focus version contained a target sentence with main stress on the subject, and the target sentences of all fillers in the focus version had main stress on the object. As the experiment was carried out in a classroom setting, there were no different lists in which the order of the items was varied. The order of items and fillers was the same in both versions. There were always at least two fillers in between two test items.

4.3 Participants and procedure

Two groups of undergraduate students of Utrecht University participated in the experiment. All were native speakers of Dutch and none had any prior knowledge of the topic. All were taking the same course, which was unrelated to the topic of the experiment. 25 students at the same time participated in the non-focus version of the experiment. They all filled out an answer sheet while the items were projected on a big screen and played through loudspeakers. Simultaneously, in another room 20 students completed the focus version of the experiment, in the same setup. The stories were presented in written form on a big screen. A recording of the story was played to make sure everyone got the information. After the story, the

experimenter played the pre-recorded target sentence twice, while the story remained on the screen. The target sentence was not presented on the screen. After the second time the sentence was played, participants got 10-20 seconds to mark their answer on the answer sheet, by encircling 'true' or 'false' at the item number. Participants were instructed that there were no 'right' or 'wrong' answers, and that they should follow their intuition. Also, they were reminded that they should not talk to each other or exchange answers during the experiment. The experiment took around 10 minutes.

4.4 Results

No participants were excluded from the analysis. The results of Experiment 3 are given in Table 7:

Table 7: SI-rates Experiment 3

Condition 1 (focus)	Condition 2 (non-focus)
85%	55%

As in Experiment 1 and 2, more SIs were calculated when *or* was in the focus part of the sentence. This effect was observed for all items (see Appendix 3). A between-subjects analysis (over participants) revealed a significant difference between the groups: Mann-Whitney $U = 156.50$, $z = -2.436$, $p = 0.008$ (one-tailed), effect size $r = -0.36$. So also without an explicit question we observed the effect of focus on SIs.

4.5 Discussion

The results of Experiment 3 show that SIs are sensitive to focus, also if the target sentence is not preceded by an explicit question but the focus structure of the target sentence is reflected by the stress pattern of the sentence. We can conclude the difference in SI-rates between the non-focus and the focus condition in Experiments 1 and 2 is not due to the verbatim repetition of *or* from the explicit question in the non-focus condition. Experiment 3 showed that when this repetition was not present, participants still calculated less SIs in the non-focus condition.

If we compare the SI-rates of Experiment 3 to the total SI-rates of Experiments 1 and 2, we see that the observed percentages were similar, but slightly higher. The 85% in the focus condition was the highest percentage observed so far and approached ceiling level (16 out of 20 participants

always calculated the SI). The difference between the two conditions was also slightly bigger (30% in Experiment 3 versus 18% and 26% in Experiments 1 and 2), supporting the view that there might have been carry-over between the conditions in Experiments 1 and 2.

However, even after controlling for carry-over effects the observed difference was not as big as predicted by the theory. In the next section I will consider some possible explanations for this. For now, I conclude that the fact that the effect is replicated with spoken stimuli and implicit questions is strong support for the focus-sensitivity of SIs.

5. General Discussion Experiments 1-3

The results of Experiment 1-3 are summarized in Table 8, where the first item data of Experiments 1 and 2 are added between parentheses.

Table 8: Summary SI-rates Exp 1-3

	focus	nonfocus
Exp 1	73% (71%)	55% (37%)
Exp 2	67% (61%)	41% (26%)
Exp 3	85%	55%

Despite the differences between the experiments (with and without contextual manipulation, direct and indirect speech, with and without explicit question, within- and between-subjects, written and spoken materials) we see a robust effect of focus on SIs, in the direction predicted by the focus-sensitive theories of Van Kuppevelt and Van Rooij. We can conclude information focus is indeed a contextual property that affects SIs.

However, the differences between the two conditions was smaller than predicted by the theory. Although the difference becomes bigger if we consider the first item data in Experiments 1 and 2, it is far from a 0-100 contrast. This might be due to the fact that SIs are hard to test experimentally, as I explained in the previous chapter and was illustrated by Chierchia et al.'s (in press) study discussed in section 2. However, in this section I will explore three other possibilities which I find plausible: the possibility of chance performance in the non-focus conditions (5.1), a possible clash between the implicit and the explicit question (5.2), and the possible effect of a default focus position (5.3).

5.1 Chance performance in the non-focus conditions: focus as a 'cue' for SI-calculation?

So far, I have only compared the two conditions without looking at the behavior of participants within one condition. However, the percentages of the non-focus conditions, resp. 55%, 41% and 55%, are close to 50%. This could be indicative of guessing behavior in this condition. Indeed, if we conduct a binomial test on the true/false data of Experiment 1, the 59 'false' versus 49 'true' answers in the non-focus condition do not differ significantly from chance ($p = 0.387$ (two-tailed)). But as this was a repeated-measures design, we should look at the distribution of participants over the possible types of behavior. There were three items per condition so there were four possible types of behavior (3, 2, 1 and 0 'false' answers (SIs) out of 3). The distribution of participants in the non-focus condition of Experiment 1 is given in Table 9:

Table 9: distribution of participants over behavior non-foc. cond. Exp. 1

behavior	freq. (parts.)
3/3 'false'	12
2/3 'false'	7
1/3 'false'	9
0/3 'false'	8

If participants were guessing we would expect both the 2/3 and 1/3 'false' behavior to be three times as frequent as the 3/3 and 0/3 'false' behavior (as there are three possibilities for 2/3 and 1/3, and only one for 3/3 and 0/3). So with 36 participants, we would expect the distribution to look like Table 10:

Table 10: expected distribution of participants over behavior if guessing

behavior	freq. (parts.)
3/3 'false'	4.5
2/3 'false'	13.5
1/3 'false'	13.5
0/3 'false'	4.5

A chi-square test reveals the observed distribution is significantly different from this expected distribution: $\chi^2 = 19.85$, $p < 0.001$. Therefore, we can conclude participants were not guessing in the non-focus condition of Experiment 1.

In the non-focus condition of Experiment 2, the ratio of 56 'false' answers versus 82 'true' was significantly different from chance on a binomial test: $p = 0.033$. The participants distribution, given in Table 11, was also significantly different from the expected values based on guessing behavior, also given in Table 11: $\chi^2 = 40.49$, $p < 0.001$. Again, we can conclude the SI-rate of around 50% (41%) is not due to guessing.

Table 11: distribution of participants over behavior n-foc. cond. Exp. 2

behavior	expected (guessing)	observed
3/3 'false'	5.75	11
2/3 'false'	17.25	6
1/3 'false'	17.25	11
0/3 'false'	5.75	18

In the non-focus version of Experiment 3, the 41 'false' vs. 34 'true' answers were not significantly different from chance: $p = 0.489$. Nevertheless, the distribution of participant behavior, given in Table 12, was significantly different from the expected values based on guessing: $\chi^2 = 9.40$, $p = 0.024$. So again, it seems participants were not guessing.

Table 12: distribution of participants over behavior non-focus version Exp. 3

behavior	expected (guessing)	observed
3/3 'false'	3.125	11
2/3 'false'	9.375	3
1/3 'false'	9.375	2
0/3 'false'	3.125	9

However, the distribution in Table 12 shows an interesting pattern: it is clearly bimodal, with hardly any participants 'in the middle'. This could indicate another type of guessing behavior, where a participant makes a guess for the first item, and sticks to this choice throughout the experiment. The fact that this pattern is more pronounced in Experiment 3 than in Experiments 1 and 2 could be due to the fact that there were no interfering focus-items in Experiment 3 (because of the between-subjects design), or that the spoken stimuli blocked alternative focus structures of the target sentence which could be superimposed on the written target sentences in Experiments 1 and 2. If the behavior in the non-focus conditions is indeed

due to an initial guess, an alternative explanation for the focus-sensitivity is possible: the hypothesis that focus is not a necessary condition for SI-calculation, but merely a 'cue'. In absence of the cue participants behave at chance (but stick to their initial choice), while if the cue is present participants will more likely than chance calculate the SI. To see whether this hypothesis is viable, we have to reconsider the first item data in Tables 5 and 6 in section 3.5, repeated here.

Table 5: Percentage SIs on all items and first items per condition Exp. 1

	focus	non-focus
all items	73%	55%
first item of that condition	75%	50%
only very first test item	71%	37%

Table 6: Percentage SIs on all items and first items per condition Exp. 2

	focus	non-focus
all items	67%	41%
first item of that condition	63%	35%
only very first test item	61%	26%

If participants were guessing in the non-focus conditions and stuck to their initial guess, we should observe first item percentages for this condition of around 50%. While this was the case for the first non-focus items of Experiment 1, it seems to not have been the case in Experiment 2, where only 35% of the first non-focus items were judged 'false'. A binomial test shows this is (marginally) significantly different from chance ($p = 0.054$). If we consider only those first items of the non-focus condition which were the first test items the participant encountered, we would also expect a percentage around 50% on the hypothesis that participants were consistently pursuing an initial guess. However, this was again not the case for Experiment 2, where the 26% is significantly different from chance ($p = 0.035$). I conclude that the hypothesis that participants made an initial guess in the non-focus conditions is possible for Experiments 1 and 3, but unlikely for Experiment 2. This makes the hypothesis of focus as a cue for SI-calculation less attractive. I now turn to the second possible explanation for the fact that the difference was smaller than predicted: a mismatch between implicit and explicit questions in the experimental items.

5.2 Inconsistency between implicit question and explicit question

Both Van Kuppevelt and Van Rooij claim the focus structure of a sentence can also be triggered by an implicit question in the context. However, it is not clear what happens when a context triggers an implicit question, but another question is asked explicitly. Therefore, a possible explanation for the unexpected high SI-rate in the non-focus conditions could be that although the scalar term appeared in the background relative to the explicit question, relative to the implicit question it was part of the information focus. To see how this could have been the case, reconsider the example of the non-focus condition of Experiment 1, represented in a table for expository reasons.

Table 13. Example (4) of the non-focus condition of Exp. 1

Katja and Birgit went searching for marine animals on the beach at their grandparents' house.	intro
Their grandfather had told them to look for a crab or a starfish. He had promised them that the one who would find a crab or a starfish, would get ten bucks.	reason question
After some searching Katja found a crab. Not much later she also found a starfish. Birgit couldn't find anything and had to return to the house empty handed.	situation
A: "Who found a crab or a starfish?"	expl. question
B: "Katja found a crab or a starfish."	target

By including the material in the second row, I provided a reason why speaker A would ask the question he eventually asked. Therefore, at that point (after reading the second row), the implicit question *Who found a crab or a starfish?* was triggered. Due to the TVJT, the story had to provide the information that Katja found both a crab and a starfish, otherwise the SI could not be assessed. However, by describing that Katja found both a crab and a starfish and Birgit found nothing, the third row of the story answered the (pair-list) question *Who found what?* instead of *Who found a crab or a starfish?* An answer to the latter question would have been for example *Katja found a crab or a starfish, but Birgit didn't*, but that description of the situation was impossible as the target sentence would always be true, irrespective of the SI. So the situation that is given necessarily answers the question *Who found what?* Relative to this question, both the subject and the direct object are information focus. It is therefore not unlikely that even though the explicit question that was asked by speaker A was only about the subject, the direct object was still considered questioned. For that reason participants might have calculated the SI, even though *or* was in the background relative to the explicit question. This possibility holds for all

three experiments, as due to the TVJT the story always had to contain the information that one person did A and B. In the next chapter I switch to another paradigm in order to avoid this problem. The third possible explanation for why the difference is smaller than predicted is related to this one, as it also raises the issue whether explicit focus marking can overrule a certain focus expectation.

5.3 Default position of information focus at the end of the sentence

The third possible explanation for the fact that the difference between the conditions was smaller than predicted by the theory is that in the non-focus conditions the focus manipulation was competing with the default focus position. Languages like English and Dutch tend to express old information at the beginning of the sentence, and new information at the end of the sentence (Halliday's 1967 'theme-rheme' structure). Therefore, hearers expect the focus part of the sentence to appear at the end of the sentence. In the non-focus conditions I tried to overrule this preference by explicitly triggering focus on the subject at the beginning of the sentence, through an explicit question in Experiments 1 and 2 and through stress in Experiment 3. However, results from previous experimental studies on the effect of focus on ambiguous sentences indicate that the default cannot be completely overridden.

Carlson et al. (2009), building on the work of Frazier and Clifton (1998), tested the preferred interpretation of ambiguous sluicing sentences like (15):

(15) The lawyer insulted the witness, but I don't remember who else.

The sentence is ambiguous, as *who else* can be short for *who else the lawyer insulted*, or for *who else insulted the witness*. On the first reading, the object *the witness* is taken as the antecedent for *who else*, and on the second reading the subject *the lawyer* is the antecedent. Carlson et al. showed with a questionnaire that in 72% of the cases, the object was the preferred antecedent. They concluded that since they controlled for the effect of recency (the distance between *who else* and the candidate antecedents) and lexical bias (they included a condition in which the same lexical items were used in switched order), the object preference is probably due to the fact that participants expect the object to be the focus of the sentence, and therefore a more likely antecedent.

Interestingly, Carlson et al. also conducted an experiment in which they manipulated the main stress of the sentence which contained the antecedent. They compared sentences like (16) and (17):

(16) The CAPTAIN talked with the co-pilot, but we couldn't find out who else.

(17) The captain talked with the CO-PILOT, but we couldn't find out who else.

Here, the main stress of the sentence indicates which constituent is the focus: the subject *captain* in (16) and the object *co-pilot* in (17). Based on this we expect a great preference for the subject to be the preferred antecedent in (16), and the object in (17). As expected, in (17) in 88% of the cases the object was picked as the antecedent, but in (16) only in 58% of the cases the subject was the preferred antecedent. This shows that the initial preference for an object antecedent is lowered by the stress on the subject in (16), but in 42% of the cases participants still preferred the object antecedent, despite the stress manipulation. Carlson et al. conclude that '[...] listeners retained a bias toward interpreting the last argument as the focus of the first clause, even when overt focus markers did not support that preference.' (p. 15).

This might have also been the case in Experiments 1-3 above. Even though in the non-focus conditions the explicit question or the sentence stress made it clear that the subject was the focus, participants might have still interpreted the direct object as the focus, in line with the default focus-position. This would explain the SI-rates of resp. 55%, 41% and 55% in Experiments 1-3, which are in the same range as the 42% found by Carlson et al. I will return to this possible explanation in chapter 7, where I test a scalar term in subject position instead of direct object position.

In the next section I evaluate the experimental paradigm that was adopted in Experiments 1-3: the TVJT. I first discuss its suitability for testing the specific hypothesis of this work, followed by a critical evaluation of the suitability of the TVJT to assess SIs in general.

6. The suitability of the TVJT

In 5.2 above we saw how the TVJT might have been responsible for the difference between the two conditions being smaller than predicted. As the TVJT required the situation to be given, it triggered a possible clash between the implicit question and the explicit question. In 6.1 I will consider how on the other hand the use of the TVJT gives room for an alternative explanation of the observed effect. These problems show that the TVJT is not an optimal experimental paradigm to test the relation between focus and SIs. Moreover, in 6.2 I will argue that the SI is not very suitable to test SIs in general, due to

the requirement that the actual situation is given. This will result in the use of a new experimental paradigm in the next chapter.

6.1 The suitability of the TVJT to assess focus-sensitivity of SIs

A disadvantage of the TVJT is that it asks for a judgment about a sentence instead of an interpretation. This opens the road for an alternative explanation of the effect observed in Experiments 1-3. Reconsider the non-focus conditions of Experiments 1-3. I repeat the non-focus condition of Experiment 1 here for easy reference:

(4) Experiment 1 Condition 2: non-focus
Katja and Birgit were searching for marine animals on the beach near their grandparents' house. Their grandfather had encouraged them both to go look for a crab or a starfish. He had promised them that the one who would find a crab or a starfish, would get ten bucks. After some searching Katja found a crab. Not much later she also found a starfish. Birgit couldn't find anything and had to return to the house empty handed.

A: "Who found a crab or a starfish?"
B: "Katja found a crab or a starfish."

Is the answer of speaker B true or false? true / false

A participant who calculated the SI and interpreted the target sentence as *Katja found a crab or a starfish but not both*, was faced with a dilemma: On the one hand, the SI makes the sentence false. On the other hand, assuming the SI does not arise in the question, it *was* Katja who found a crab and/or a starfish, not Birgit. Therefore, the part of the answer that Katja was the one who found a crab and/or a starfish is true. The participant might have reasoned as follows: 'Although the answer stated this way means/implicates that Katja found one of the two, the answerer has provided the correct information that it was Katja, and not Birgit, who found a crab and/or a starfish.' This might have persuaded participants to judge the sentence true, even though they calculated the SI. In the focus conditions there was no such dilemma as the SI related directly to the question what Katja brought. Therefore, this is an alternative explanation of the difference between the conditions in Experiments 1-2. The explanation extends to Experiment 3 as relative to the implicit question the answer in the non-focus condition was also partly true.

The problem is caused by the fact that the TVJT asks for a judgment about the sentence, instead of directly assessing the interpretation of the sentence. In the next chapter I switch to a new paradigm which asks directly about the interpretation of the target sentence. However, the disadvantages of the TVJT are not limited to testing our specific hypothesis of SIs and focus. In the next section I will argue that the SI is a very artificial way to assess SIs in general.

6.2 The suitability of the TVJT to assess SIs in general

Consider how an SI arises in everyday language: A speaker utters a sentence with a scalar term about something that happened. The addressee, who did not already know what happened (we are assuming the speaker's utterance was informative), interprets the sentence and draws an inference based on the use of the scalar term by the speaker. She concludes that the situation which would be described by the stronger scalar term is not what happened. Finally, she updates her knowledge with the complete meaning of the sentence (including the SI).

Now compare this to the task a participant is given in the TVJT: The addressee (in this case the participant) is presented with a situation. Therefore, she knows exactly what happened. After that, a fictive speaker uses a scalar term. Now, in order for the TVJT to be a good measure of SIs, the participant should interpret the sentence and draw an inference about what happened based on the use of the scalar term. After that, she has to compare the complete meaning (the meaning of the sentence including the SI) to the situation that was provided earlier, and determine whether this complete meaning is true relative to the story. If she calculated the SI, she will judge the sentence false based on this comparison.

It is very unlikely that these two final steps (calculate the SI, compare complete meaning to story) will be taken by the participant. After all, why should the participant draw an inference about what happened if she already knows what happened? In order for the TVJT to trigger SIs, the participant has to interpret the sentence *as if she didn't know what happened*. It is very unlikely that a participant will do this. A much more straightforward way to complete the task is to check whether the target sentence is a good match with the story. However, the target sentence typically contains a weaker scalar term than could have been used. Therefore, a participant who is checking the sentence with the story will notice this discrepancy. It is therefore quite likely that she will reject the target sentence based on the fact that a better match would have been possible, the sentence with the stronger scalar term.

Notice that for this the participant does not have to go through the whole process of interpreting the target sentence as if she didn't know what happened and drawing the SI. There is no need to calculate an SI in order to see that a stronger term could have been used. For that, only knowledge of the scale is necessary. Therefore, it is much more likely that a 'false' answer in the TVJT is due to the participant considering the use of the weaker term unacceptable, than it is due to the participant calculating an SI. In Experiments 1-3 participants were instructed to ignore strangeness or ugliness of the sentence, but it is questionable whether they consider the use of a weaker term to fall under these descriptions. One could claim that if in a certain environment an addressee who knows what happened considers a weaker term unacceptable, it is likely that in the same environment an SI will be calculated by an addressee who does not know what happened. However, this relation is not a priori true, and it needs experimental support. Also, on that view the TVJT would be a rather indirect, and therefore not very accurate, measure of SIs.

The fact that 'false' answers in the TVJT do no necessarily indicate SIs, but are at least as likely due to unacceptability of a weaker term, is another explanation for why the difference between the two conditions was smaller than predicted in Experiments 1-3. A better measure for SIs would be a paradigm in which participants do not know the actual situation, and are asked to interpret a sentence with a scalar term. In the next chapter I present a number of experiments in which the focus-sensitivity of SIs was tested in a new paradigm which satisfies this requirement.

However, in the next section I first consider two alternative explanations (to the focus-sensitive SI-theories) of why focus leads to the calculation of more SIs. Both of these explanations stem from the psycholinguistics literature. Although my conclusion will be that these alternative explanations cannot account for the data of Experiments 1-3, they are important for the discussion as they are based on empirical observations of the effects of focus on language processing.

7. Alternative hypotheses from the psycholinguistic literature on why focus leads to more SIs

Throughout Experiments 1-3 I already controlled for a number of alternative explanations based on other differences between the conditions than focus (e.g. presence of *or* in a DE-environment in the story, epistemic states of the speakers, repetition of *or* from the question). However, there are also other explanations for why more SIs are observed if a scalar term is in the information focus part of the sentence than the SI-accounts of Van

Kuppevelt and Van Rooij. I will discuss two of these alternative hypotheses in this section: the hypothesis that focus leads to deeper processing (7.1) and the hypothesis that emphasis on or leads to extra affort or scale activation (7.2).

7.1 The hypothesis that focus triggers deeper processing

In the psycholinguistic literature a substantial number of studies have been reported that tested the effects of focus on processing. Many of these papers conclude that material in the focus is processed better than material in the background. Therefore, this could be an alternative explanation for why more SIs were observed in the focus conditions of Experiments 1-3. However, the notion of *focus* is used wider in the psycholinguistic literature than I have been using it so far. Where I adopted Jackendoff's definition of the (information) focus of the sentence as the part of the sentence that contains new information, Birch and Rayner (1997) define focus as follows:

> 'the focus of a sentence consists of the information that is newly asserted in a discourse, sometimes contrastive, and is most prominent or emphasized within the sentence' (p. 653).

On this definition, *focus* is similar to *emphasis*, which also explains why the verbal form (*to focus* certain material) is used often in this literature. Of course it is no coincidence that *focus* is used for this wider notion based on emphasis. There is a close connection between emphasis and information focus: new information is usually presented intonationally prominently, as the information focus of the sentence normally receives the main stress in the sentence (see chapter 2). Furthermore, other forms of emphasis, such as *it*-clefts, have also been claimed to be a syntactic expression of the grammatical focus feature.

However, the problem of collapsing the two notions is that not every time a word or constituent is emphasized, it is necessarily due to it being the information focus. Consider for instance the exchange in (18), taken from Carlson et al. (2009):

(18) A: John introduced Sally.
B: (No,) John introduced MARY.

Here, *Mary* is emphasized intonationally, to indicate a contrast with previously mentioned material. It is also the information focus, but the intonational contour of *Mary* in this case is different than if B's answer

would have been a normal information focus, e.g. when B's answer (without 'No') would have been an answer to *Who did John introduce?* (see Pierrehumbert & Hirschberg 1990). This led some scholars to introduce the notion of *contrastive focus* for sentences like (18), setting it apart from information focus (see Kiss 1998 i.a.). Sentences in which material is emphasized by an *it*-cleft have also been claimed to belong to this category. Therefore, some of the results in the psycholinguistics literature might actually not be about the effect of *information* focus on processing. Let us nevertheless look at some studies in more detail to see if they relate to our findings. In order to avoid confusion over terms, I will use the term *emfocus* (a contraction of *emphasis* and *focus*) to refer to the wider notion of focus, and *information focus* to refer to the meaning of focus I have been using so far: the new information part of the sentence.

One of the oldest studies on the processing of emfocus that is often referred to is the study by Cutler & Fodor (1979), who tested the effect of emfocus on phoneme recognition. They compared spoken question-answer pairs like (19) and (20) on how fast the phoneme /b/ was recognized by participants. I will use underlining to indicate the questioned constituent.

(19) emfocus condition:
Q: What hat was the man wearing?
A: The man on the corner was wearing the blue hat.

(20) non-emfocus condition:
Q: Which man was wearing the hat?
A: The man on the corner was wearing the blue hat.

Their results show that /b/ in *blue* was recognized faster in (19), when it was emfocus. This experiment illustrates the difference between emfocus and information focus. Even though *blue* was only questioned in (19), it was also new information relative to the question in (20). Nevertheless, as Cutler & Fodor used spoken stimuli, *blue* was probably deaccented in (20), indicating it was old information. Therefore, their results might extend to information focus anyway.

Emfocus is also claimed to facilitate lexical access. Blutner & Sommer (1988) used the same kind of manipulation as Cutler & Fodor, but used a lexical decision task for a synonym of the target word. They compared (German versions of) conditions like (21) and (22):

(21) emfocus condition:
Q: Which opening did the guests delay?
A: The guests from abroad delayed the opening of the ball.

(22) non-emfocus condition:
Q: Which guests delayed the opening?
A: The guests from abroad delayed the opening of the ball.

A synonym of *ball* was recognized faster in (21) than in (22). Although similar to Cutler & Fodor's study, *ball* was new information in both conditions and therefore extending the results to information focus might be tricky, it can at least be concluded that emfocus facilitates lexical access.

The effect of emfocus on processing was also observed for so-called *semantic illusions*. These semantic illusions were studied by Erickson & Mattson (1981), who presented participants with questions in which a wrong but related name was used, e.g. *How many animals of each kind did Moses take on the Ark?* Many participants didn't notice the use of *Moses* instead of *Noah* and answered 'two'. This famous example is often called the *Moses Illusion*. Bredart & Modolo (1988) showed that this effect is sensitive to emfocus triggered by *it*-clefts. They compared written conditions like (23) and (24):

(23) emfocus condition:
It was Moses who took two animals of each kind on the Ark.

(24) non-emfocus condition:
It was two animals of each kind that Moses took on the Ark.

In (23), the anomaly is noticed more often than in (24). This suggests that the lexical material in emfocus is processed deeper or more thoroughly than in non-emfocus. Again, due to the lack of prior context and the use of an *it*-cleft as focusing mechanism, it is unclear whether these results also hold for information focus.

A number of studies have also addressed the effects of emfocus on reading times, e.g. Birch & Rayner (1997), Morris & Folk (1998), Ward & Sturt (2007). The results of these studies are not converging so I will not discuss them here. I will return to these studies in chapter 6, when I discuss the reading time experiments.

Finally, emfocus has been claimed to affect change detection. Sturt et al. (2004) tested written conditions like (25) and (26):

(25) emfocus condition:
Everyone had a good time at the pub. A group of friends had met up there for a stag night. What Jamie really liked was the cider, apparently.

(26) non-emfocus condition:
Everyone had a good time at the pub. A group of friends had met up there for a stag night. It was Jamie who really liked the cider, apparently.

When a second screen was presented in which the word *cider* was changed, it was noticed more often in (25), where it was emfocus due to a cleft, than in (26), where the cleft emphasized *Jamie*. This result is also hard to extend to information focus because in both conditions *cider* is new information.

Summing up, results from psycholinguistics have shown that emfocus facilitates phoneme recognition, lexical access, detection of semantic anomalies and change detection. This has led some authors to claim that emfocused material is processed deeper or more thoroughly than non-emfocused material (see e.g. Sanford & Sturt 2002), and that non-emfocused material is processed 'shallowly', or represented in 'good enough representations' (Ferreira et al. 2002).

Even though the results of the experiments on emfocus discussed above cannot straightforwardly be extended to information focus, let us still explore the hypothesis that our results are due to this kind of shallow processing in the non-focus conditions. There is one important reason why I think the effect of shallow vs. deep processing does not apply to our experiments, and that is that the test items and the task of Experiments 1-3 were very explicit. The observations on the effects of emfocus above were all about processing: emfocused material is processed faster and better, and changes are detected more often. However, Experiments 1-3 were off-line tasks in which there was no time pressure (perhaps some in Experiment 3). Furthermore, as the target sentences of all the items and the majority of the fillers contained *or*, it is likely that participants have noticed this and therefore have been paying attention to when *or* was used. Hence, it seems very unlikely that participants due to shallow processing did not notice that *or* was used in the items of the non-focus conditions. Moreover, as I argued above when I was discussing the carry-over effect, the test items clearly stood out, because of the discrepancy between the *A and B* story and the *A or B* sentence. It is therefore unlikely that participants did not (or less often) notice this discrepancy in the non-focus conditions. In the next section I discuss another alternative hypothesis from the psycholinguistics literature on why more SIs are triggered in focus. This hypothesis is directly about the relation between SIs and (em)focus.

7.2 The hypothesis that emphasis on *or* itself leads to extra effort or scale activation

Other psycholinguistic studies have investigated whether emphasizing only *or* itself (instead of the whole constituent as in Experiment 3) leads to more SIs. In this section I discuss two of these studies, and consider whether their explanation of the data can be an alternative explanation for the effect found in Experiments 1-3.

The first study is a study by Chevallier, Noveck, Nazir, Bott, Lanzetti & Sperber (2008).[10] They presented participants with five letters on a screen (both words and nonwords were included), for instance the letters T, A, B, L and E, followed by a sentence of the form *There is an A or an B*, see (27):

(27) T A B L E
There is an A or a B.

Participants were asked to judge the sentence with reference to the five letter string. A 'false' answer indicated participants calculated the SI (as for instance TABLE contains both an A and a B). The critical manipulation was that *or* was either emphasized or not. This was done in two modalities: written, by capitalizing and underlining *or* and oral, by stressing *or*. The four conditions are represented in (28), where italicized capitals mark prosodic stress.

(28) C1: written unstressed: There is an A or a B.
C2: written stressed: There is an A <u>OR</u> a B.
C3: spoken unstressed: "There is an A or a B."
C4: spoken stressed: "There is an A *OR* a B."

Chevallier et al. tested these conditions in a between-subjects design (one condition per group). The results were as follows: C1: 19% SIs, C2: 42% SIs, C3: 23% SIs, C4: 73% SIs.[11] So both in written and spoken form, the SI-rate goes up as a result of the emphasis on *or* itself.

[10] I thank Ira Noveck for pointing out the existence of this paper to me, which I had somehow overlooked myself.

[11] As I mentioned before, in a within-subjects follow-up study in which only C3 and C4 were compared, the difference was much smaller: C3: 32%, C4: 52%. This once again points to the effect of carry-over.

Chevallier et al. predicted this effect from a relevance theoretic viewpoint, in which there is a pay-off between effort and effect. Stressing *or* motivates the hearer to make a bigger effort in interpreting the sentence. Therefore, the effect should also be bigger and it is more likely the hearer will go beyond the literal meaning of the sentence and enrich the meaning by drawing the SI.[12] One could argue that the reason why more SIs were observed in the focus conditions in Experiments 1-3 is this increased effort triggered by a marked form. This explanation seems to match especially well with Experiment 3, in which *or* was part of the constituent *A or B* that received the main sentence stress. However, I argue the situation in which only *or* is emphasized is not comparable to the situation I tested, in which the whole constituent is the information focus.

The sentences of Chevallier et al. in which only *or* is emphasized give rise to a *contrastive focus*: they can only be used to indicate a contrast with the sentence in which the stronger scalar term is used: *There is an A and a B*. There is no question to which these sentences (with this stress pattern) are a natural answer. Therefore, it is not surprising that the rate of exclusive readings goes up once *or* is emphasized: the sentence can only be a correction of, or a contrast with the *and*-sentence, so the fact that the *and*-situation is not the case follows naturally.[13] The sentences I tested in Experiments 1-3 do not have a marked stress pattern. In Experiments 1 and 2 the sentences were written without any marking of stress, so no extra effort caused by a marked form could have occurred there. Even if participants silently attributed a stress pattern to the sentences, this would be the natural pattern for the answer to the question. Hence, the explanation of Chevallier at al. of their data cannot explain the observed effect in Experiments 1 and 2.

The items in Experiment 3, of which I repeat the example of the focus condition here for convenience, did contain a stress pattern that was different from the neutral stress pattern of the sentence:

(11) Experiment 3 target sentence focus condition:
Paola took AN APPLE OR A PEAR from the fruit section.

[12] This strikes me as very similar to the M-principle of Levinson (2000) that marked expressions have marked meanings. See also Krifka (2002) for the Bidirectional OT version of this principle (superoptimality).

[13] It is questionable whether participants actually calculated an SI here, as the *and*-situation is already ruled out by the contrast. In that sense this situation is comparable to a disjunction with mutually exclusive disjuncts, such as *John is in Paris or in Rome*. Here too, the question is whether the SI is calculated anyway or whether it is not calculated because it is not needed.

However, the stress pattern in (11) is not marked, it is the natural stress pattern for an answer to the question about the direct object, which was triggered by the context. If any extra effort would be attributed to interpreting the sentence due to the main stress on the whole constituent *an apple or a pear*, it would be taken to indicate that this part of the sentence is the information focus (so it answers the question about the direct object), which is exactly what we wanted to achieve. Therefore, I conclude Chevallier et al.'s explanation can also not be extended to the data of Experiment 3.

Another group of researchers that investigated the effect of stress on *or* itself are Schwarz, Clifton & Frazier (in progress). They claim that emphasizing a scalar term increases SI-rates because it *activates the scale*:

> 'It is commonly noted, at least in passing, that focusing a scalar item increases a listener or reader's tendency to compute a scalar implicature, presumably because focus draws attention to the speaker's use of a particular term on a scale and thus activates the scale itself. By activating the scale, the contrast between the term used and its alternatives is highlighted.' (p. 7)

They tested this hypothesis in a paraphrase selection study, in which spoken sentences like (29) and (30) were compared. In (29) *or* was emphasized, similar to the items of Chevallier et al. above, and in (30) the auxiliary *will*.

(29) Mary will invite Fred OR Sam to the barbecue.

(30) Mary WILL invite Fred or Sam to the barbecue.

In a between-subjects design, sentences like (29) returned 84% SIs and sentences like (30) 71%.[14] Again, stressing *or* gives rise to a contrastive

[14] The reason these rates (especially in the non-stressed condition) are much higher than those of Chevallier et al. might be due to the difference in the sentences (real-life situations versus abstract statements), or the difference in task (paraphrase selection versus truth value), but I think one more difference might be important: Chevallier et al. used existential sentences (*There is an A or a B*). Firstly, these were clearly not exhaustive: only two letters were mentioned while 5 were presented. Therefore, the sentences could have not been an answer to the mention-all question *What letters are there on the screen?* It is well-known that answers to mention-some questions are less likely to trigger SIs (see Van Rooij 2002). Secondly, existential sentences might in general trigger fewer SIs due to the scope interaction of the quantifier and the disjunction (consider *There is something which is an A or a B*, which is true in Chevallier at al's items irrespective of the reading of *or*).

focus. Schwarz and colleagues confirm this by indicating that a L+H* accent was used on *or*, which according to Pierrehumbert & Hirschberg (1990) is used to mark a correction or contrast. So indeed, if *or* itself is emphasized the alternative with *and* is activated, for the simple reason that it is the only alternative for which a sentence like (29) could be a correction or contrast. However, the question is whether this observation using (contrastive) emphasis on *or* can be extended to the case where a whole constituent of the form *A or B* is the information focus, so whether our results in Experiments 1-3 can be due to the scale being activated in the focus case and not in the non-focus case. It seems a lot hinges on what it means for a scale to be *activated*. If it just means that the alternatives are considered, it is unlikely that the results of Experiments 1-3 are due to this. Above I already mentioned it is unlikely that participants did not notice the discrepancy between the *A and B* story and the *A or B* sentence in the non-focus conditions. Therefore, it seems unlikely that they did not consider the *and*-alternative in that condition. If activation of the scale means something else than the hearer considering the alternatives, it seems the question of when the scale is activated is reduced to the question when the SI is calculated. Therefore, without a clear theory of what it means for a scale to be activated and when this happens, this does not add much.

8. Summary and Conclusions

As this was quite a long story, it is time to recapitulate. In sections 2-4 I presented Experiments 1-3 which tested the hypothesis that more SIs are calculated if a scalar term appears in the part of the sentence that is the information focus. In Experiment 1 focus was manipulated by an explicit question which was supported by the context. In Experiment 2 I brought back the two conditions to a minimal pair and only manipulated the explicit question. Experiment 3 extended the investigation to implicit questions by using spoken stimuli in which stress indicated which part of the sentence was the information focus. In all three experiments significantly more SIs were observed in the focus conditions. The results are repeated here.

Table 8: Summary SI-rates Exp 1-3

	focus	nonfocus
Exp 1	73% (71%)	55% (37%)
Exp 2	67% (61%)	41% (26%)
Exp 3	85%	55%

Although the results show a clear pattern, the difference was not as big as predicted by the theory. The first item data indicate that in Experiments 1 and 2 there might have been some carry-over from the focus conditions to the non-focus conditions, but in Experiment 3 this is out due to the between-subjects design.

In section 5 I explored three possible explanations for the fact that the difference was smaller than predicted. The first one was that participants might have been guessing in the non-focus conditions, as the SI-rates are around 50%. However, the distribution of the participants over the possible types of behavior was different than predicted by guessing behavior. A pattern in which participants take a guess at the first item and stick to it throughout the experiment is possible for the non-focus conditions of Experiments 1 and 3, but did not match the data of the non-focus condition of Experiment 2. The second explanation for why the difference was smaller than predicted was that in the non-focus conditions there might have been a mismatch between the explicit question and the implicit contextual question. As the TVJT required that the actual situation was given, this might have triggered an implicit *Who found what?* question, causing the direct object to be part of the information focus of the target sentence in the non-focus conditions too. This is one of the problems I will avoid by switching to another experimental paradigm in the next chapter. The third possible explanation for the smaller difference is that the default position of the information focus is at the end of the sentence. Results from the literature showed that explicit marking of focus by stress can only partly overrule participants' preference for a sentence-final focus. If despite the question and stress-manipulations participants in a reasonable number of cases still considered the sentence-final direct object to be focus in the non-focus conditions, this could explain why we still find a reasonable number of SIs in these conditions. I come back to this explanation in chapter 7, where I test a scalar term in subject position.

In section 6 I discussed the suitability of the experimental paradigm that was used, the TVJT. The first disadvantage of the paradigm is that it opened the road for an alternative explanation of the effect. One could argue that in the non-focus conditions the target sentence with SI was still partially true, as it provided the correct information which person in the story satisfied the questioned predicate. As this dilemma did not arise in the focus conditions, it might explain the observed effect. This problem is the result of the TVJT asking for a judgment about a sentence instead of the interpretation. The new paradigm I present in the next chapter does assess interpretation directly. However, I argued the TVJT is not just problematic to test the hypothesis of SIs and focus, but that it is a very indirect measure of SIs in general. As the participant already knows what happened, it does not provide a natural

environment for SIs to arise. Participants do not have to go through the process of calculating an SI in order to reject the target sentence, they only have to notice the possibility of a stronger term. In the new paradigm, presented in the next chapter, the actual situation is taken out so participants do not know what actually happened.

I presented two alternative explanations of the effect of focus on SIs in section 7. The first was based on experimental results in the literature which showed the effects of focus on phoneme recognition, lexical access, detection of semantic anomalies and change detection. This led some authors to propose that focus triggers deeper processing, which could explain why more SIs were observed in the focus conditions. However, while the studies referred to showed advantages in speed and detection in on-line tasks, Experiments 1-3 were off-line tasks which were very explicit and in which there was no time pressure. It is therefore unlikely that participants due to shallow processing did not notice the presence of *or* in the non-focus conditions. The second alternative explanation for the effect of focus on SIs was based on observations in experiments in which only *or* itself was stressed. I claim this situation is not comparable to the situation I tested, as stressing *or* itself gives rise to a contrastive focus instead of a mere information focus. It is not surprising that stressing only *or* itself leads to many SIs, as the sentence can only be intended as a correction of contrast with the *and*-sentence. Chevallier et al. (2008) explain this effect by claiming stress on *or* itself triggers extra effort in interpretation, while Schwarz et al. claim it activates the scale. I argue no extra effort is expected in Experiments 1-2 as no stress marking was present in the written stimuli, and in Experiment 3 the main stress of the sentence was not marked, but the natural stress contour for the answer. Therefore, the explanation of Chevallier et al. of their data does not extend to the current experiments. I also argued that if scale activation means that participants considered the stronger alternative, this cannot account for the effects in Experiments 1-3, as it is unlikely that participants did not consider the *and* alternative in the non-focus conditions. Therefore, the explanation of Schwarz et al. also does not provide a good alternative explanation for our data.

I conclude that the results support the hypothesis that more SIs arise if a scalar term is part of the information focus of the sentence. However, from a methodological point of view the TVJT has a number of serious shortcomings for assessing SIs. Therefore, in the next chapter I will assess the effect of focus on SIs in a new paradigm which avoids the shortcomings of the TVJT. I will also try to tease apart focus and another contextual property that might have added to the effect: the relevance of the stronger scalar alternative. I will explain and test this hypothesis in the next chapter.

CHAPTER 4: TEASING APART FOCUS AND ALTERNATIVE RELEVANCE WITH THE PWJT-PARADIGM

1. Introduction

The results of the experiments presented in chapter 3 supported the hypothesis that SIs are sensitive to the contextual property of focus. However, I argued that the TVJT is not very suitable to test this hypothesis, or to test SIs in general. Therefore, I conducted four experiments (Experiments 4-7) in which I tested the focus-sensitivity of SIs in a new experimental paradigm. The goal of these experiments was not just to find additional evidence for the effect of focus on SIs, but also to tease apart the effects of focus and effects of another contextual property, the relevance of the stronger scalar alternative. In section 2.1 I present this property and I discuss how it may have contributed to the result that more SIs were calculated in the focus conditions in Experiments 1-3. In 2.2 I consider how it can be teased apart from focus in experimental conditions. The new story type that is needed for this is presented in 2.3. I present the new experimental paradigm, the *Possible World Judgment Task* (PWJT) in section 3.

The outline of the rest of the chapter is as follows: In section 4 I present the first PWJT-experiment (Experiment 4). In the discussion of this experiment I will point out three possible sources of experimental noise in the experiment. Experiment 5, presented in section 5, is a redo of Experiment 4, in which these three factors were controlled for. Experiments 4 and 5 were both computer-based experiments: Experiment 4 was a web-based questionnaire and Experiment 5 was conducted on computers in the lab. In section 6 I present two paper-and-pencil versions of the same experiments (Experiments 6 and 7). I end this chapter with a general discussion of all four experiments (section 7). There, I will motivate the switch to on-line experiments which is made in the next chapter.

2. Another player enters the field: alternative relevance

2.1 Alternative relevance

The experiments presented in the previous chapter showed that more SIs were calculated if the scalar term was part of the information focus. This was predicted by the focus-sensitive accounts of Van Kuppevelt and Van Rooij.

However, there is an alternative explanation of the data based on the Gricean nonce derivation presented in chapter 1, repeated here for convenience:

(1) i. The speaker used the scalar term *or*.
ii. The speaker could have uttered the same sentence with the scalar term *and* instead of *or*, which would have been stronger / more informative (because the sentence with *and* entails the sentence with *or*).
iii. The sentence with the stronger scalar term *and* would have also been relevant.
iv. The speaker is trying to be as informative as possible (she is obeying the Maxim of Quantity).
v. Apparently, the speaker does not have evidence for the sentence with *and*.
vi. The speaker is well informed.
vii. Therefore, it is likely that the speaker considers the sentence with *and* to be untrue.

The crucial step is step (iii), where the relevance of the stronger scalar alternative is assumed. We could explain the effect in Experiments 1-3 by reasoning about whether the stronger alternative was of interest to the hearer, in this case the questioner. Take for instance the example stories of Experiment 2, repeated here for convenience:

(2) Experiment 2 focus condition
Julie and Karin were searching for marine animals on the beach. After some searching Julie found a crab. Not much later she also found a starfish. Unfortunately, Karin didn't find anything.
When Karin returned, her mother asked what kind of marine animals Julie had found.
Karin answered that Julie had found a crab or a starfish.

(3) Experiment 2 non-focus condition
Julie and Karin were searching for marine animals on the beach. After some searching Julie found a crab. Not much later she also found a starfish. Unfortunately, Karin didn't find anything.
When they returned, their mother asked who had found a crab or a starfish.
Karin answered that Julie had found a crab or a starfish.

In (2) and (3) it is unclear why the mother asked the question. The story does not include a reason why this question was relevant. However, participants

might have accommodated assumptions about what the mother was interested in, based on the question she asked. In (2), the mother asked what Julie had found. Therefore, she was probably interested in *what exactly* Julie had found. Hence, the scalar alternative with *and* (*Julie had found a crab and a starfish*) would have also been relevant to her. In (3) however, the mother asked who caught a crab or a starfish. Therefore, she was not necessarily interested in what exactly Julie caught, but merely in who of Julie and Karin succeeded in finding a crab or a starfish. Consequently, the stronger alternative with *and* was not necessarily also relevant to her. As the nonce derivation requires the stronger alternative to be relevant, the difference in assumptions about whether or not the stronger alternative was of interest to the hearer might have caused the observed difference in SI-rates.

In Experiment 3 the same reasoning applies. There was no explicit question, but the stress pattern of the target sentence indicated which question it was an answer to. Again, there was no indication in the story why this question was asked, so participants might have derived what was of interest to the hearer based on the accommodated question itself. Therefore, the same difference in relevance of the stronger scalar alternative to the hearer might have been responsible for the effect.

However, in Experiment 1 I manipulated the context to support the question that was asked by including a reason why the question was relevant. The example items of Experiment 1 are repeated below.

(4) Experiment 1 focus condition

Katja was searching for marine animals on the beach near her grandparents' house. She had promised her grandfather to find some beautiful animals. He had said that if she would find an oyster, she would get ten bucks. Katja soon found a crab. Not much later she also found a starfish. But no matter how hard she looked, she didn't find an oyster.

A: "What did Katja find?"
B: "Katja found a crab or a starfish."

(5) Experiment 1 non-focus condition
Katja and Birgit were searching for marine animals on the beach near their grandparents' house. Their grandfather had encouraged them both to go look for a crab or a starfish. He had promised them that the one who would find a crab or a starfish, would get ten bucks. After some searching Katja found a crab. Not much later she also found a starfish. Birgit couldn't find anything and had to return to the house empty handed.

A: "Who found a crab or a starfish?"
B: "Katja found a crab or a starfish."

In (4), a reason to ask the question that speaker A asked is to find out whether Katja found an oyster and got the reward of ten bucks. The question asked by A in (5) could be asked to find out who of Katja and Birgit got the reward. At first glance, in both (4) and (5) it seems the stronger alternative with *and* was not relevant to the questioner, as in both conditions it was not relevant for the reward. However, let us look at the stories in more detail. In (5) it is obvious that A asked the question in reference to the reward. Why else would she specifically inquire about a crab or a starfish? However, in (4) this is not so obvious. It is not even clear whether speaker A knew about the reward and the requirements for it. If the questioner knew this and wanted to know whether Katja got the reward, she could have asked *Did Katja find an oyster?* (or just *Did Katja get ten bucks?*). Perhaps speaker A only knew Katja went out searching for marine animals and she just wanted to know what Katja found. Another possibility is that speaker A did know about the reward but was nevertheless interested in what Katja found, not just in whether she found an oyster. Therefore, participants might have considered the stronger scalar alternative to be relevant to the hearer in the focus condition but not in the non-focus condition. This could explain why more SIs were observed in the focus condition. I will call the contextual property that the stronger scalar alternative was relevant *alternative relevance*. In the next section I consider how focus and alternative relevance can be teased apart in experimental conditions.

2.2 Teasing apart focus and alternative relevance

It seems focus and alternative relevance are hard to tease apart. However, they do not necessarily co-occur. By explicitly indicating what is relevant to the hearer, we can create contexts in which focus and alternative relevance are contrasted. In (5) we already saw an example of how alternative

relevance can be explicitly denied, through the use of a conditional in which the weaker alternative (*A or B*) is the requirement. The conditional is repeated in (6): [1]

(6) He had promised them that the one who would find a crab or a starfish, would get ten bucks.

In order to get the reward, it is not relevant whether Katja found both a crab and a starfish or just one of the two. So if speaker A is asking the question in order to find out who got the reward, the stronger scalar alternative *Katja found a crab and a starfish* is not relevant to her. I will use conditionals like this in Experiment 4-7 to create contexts that lack the property alternative relevance.

How can we create contexts in which the alternative is relevant? A way to do this is to introduce a version of the conditional for which it is crucial whether Katja found both animals or just one, e.g. the conditional in (7):

(7) He had promised her that if she would find at least two marine animals, she would get ten bucks.

In (7) finding two marine animals gets Katja the reward, so it is relevant whether she found both animals. Hence, the alternative *Katja found a crab and a starfish* is relevant. This type of conditional will be used in Experiments 4-7 to trigger alternative relevance.

In order to tease apart focus and alternative relevance, we have to contrast the two properties. This can be achieved by combining the two types of conditionals in (6) and (7) with the two types of questions which manipulate focus. Consider the four possible combinations in (8)-(11):

(8) [+alternative relevance, +focus]
He had promised her that if she would find at least two marine animals, she would get ten bucks.
A: "What did Katja find?"
B: "Katja found a crab or a starfish."

[1] Strictly speaking, (6) is not a conditional. However, I will use the term *conditional* in a broad sense, as a sentence containing a requirement and the consequence of satisfying the requirement.

(9) [-alternative relevance, -focus]
He had promised them that the one who would find a crab or a starfish, would get ten bucks.
A: "Who found a crab or a starfish?"
B: "Katja found a crab or a starfish."

(10) [+alternative relevance, -focus]
He had promised her that if she would find at least two marine animals, she would get ten bucks.
A: "Who found a crab or a starfish?"
B: "Katja found a crab or a starfish."

(11) [-alternative relevance, +focus]
He had promised them that the one who would find a crab or a starfish, would get ten bucks.
A: "What did Katja find?"
B: "Katja found a crab or a starfish."

In (10) and (11) alternative relevance and focus are contrasted. However, the question in (10) does not seem to make sense after the conditional. Why would someone want to know who caught specifically a crab or a starfish, while the conditional was about finding at least two marine animals, irrespective of which type? Also, it is strange to inquire about who found at least one animal while the requirement was finding at least two animals. It seems it is hard to come up with a context in which the stronger alternative is relevant, but in which the non-focus question which contains *A or B* is asked. Fortunately, (11) is fine. If speaker A wants to find out whether Katja got the reward, this question is perfectly natural. In (11) the stronger scalar alternative *Katja found a crab and a starfish* is not relevant as it is not relevant for the reward. Nevertheless, *a crab or a starfish* is the information focus of the sentence due to the question. So we can use (11) as the critical case for teasing apart focus and alternative relevance. Therefore, I used combinations like (8), (9) and (11) in the conditions of the experiment. Assuming that the [+alternative relevance, +focus] combination in (8) will trigger more SIs than the [-alternative relevance, -focus] combination in (9), it is crucial what the [-alternative relevance, +focus] combination in (11) will do. If focus is the important property for SIs, (11) should pattern with (8) (as both are [+focus]), and trigger more SIs than (9). If alternative relevance is what is crucial, (11) should pattern with (9) (as both are [–alternative relevance]) and return less SIs than (8). A third possibility is that both focus and alternative relevance are important. In that case the SI-rate of (11) should be in between those of (8) and (9). However, to make sure the

conditional and the question are linked, we need a new story type, which I will introduce now.

2.3 A new story type

Discussing example (4) of Experiment 1 above I pointed out that it was not obvious that the questioner asked the question because of the conditional. This was due to the fact that in Experiment 1 it was unclear what the relation of the speakers was to the story. Therefore, I made a new story type in which I changed two more things from the TVJT stories. First, I introduced two speakers. One was telling the story (e.g. about Katja who went searching for marine animals), including the conditional, and the other asked the focus-determining question. This way it was clear that the questioner knew about the conditional, as she had just been told about it. Secondly, I took out the situation from the story. Compare for illustration the TVJT item of Experiment 2, repeated from (4) above, to the new story type in (12):

(4) Experiment 2 focus condition
Katja was searching for marine animals on the beach near her grandparents' house. She had promised her grandfather to find some beautiful animals. He had said that if she would find an oyster, she would get ten bucks. Katja soon found a crab. Not much later she also found a starfish. But no matter how hard she looked, she didn't find an oyster.

A: "What did Katja find?"
B: "Katja found a crab or a starfish."

(12) new story type
Marieke told her mother that Katja went searching for marine animals on the beach yesterday, and that her grandfather had told her that if she would find at least two marine animals, she would get ten bucks.
The mother said: "Oh, and what did Katja find?"
Marieke answered: "She found a crab or a starfish."

In (4) it is unclear who speaker A is, what she does and does not know about the story and why she asked the question. Apparently she knows Katja went searching for something, but she obviously does not know what the result was. So she has not been told the part of the story in which it was described what Katja found. It is therefore questionable whether she knows about the

conditional, as I already pointed out above. In (12) on the other hand, it is clear that the mother knows about the conditional, as she has just been told about it.

Another improvement from (4) to (12) is that the situation is taken out, so the question is a natural part of the story. As the outcome of the conditional is no longer already known, it is natural that the speaker inquires about this: One speaker introduces a conditional and in the next turn, the other one asks whether the requirement in the conditional was satisfied. To stress that the mother has heard the conditional and is responding to it, I included "Oh, and..." in her utterance. Together, these changes make it clear that the focus-determining question was asked because the questioner wanted to know whether the requirement of the conditional was satisfied. I will now present the new experimental paradigm, called the *Possible World Judgment Task*, which I used to test stories of the type of (12).

3. The Possible World Judgment Task (PWJT) paradigm

In the previous chapter I argued the TVJT was not a very good paradigm for testing our hypothesis. My criticism focused on two points: the requirement that the actual situation was given (e.g. the information who found what), and the fact that the TVJT asks for a judgment about a sentence instead of assessing the interpretation directly. In section 5.2 of chapter 3 I argued that the requirement that the situation was given possibly caused a mismatch between the implicit question that the context gave rise to and the (focus-determining) explicit question that was asked after the situation. In section 6.2 I proposed that the fact that the situation is given is a more general problem of the TVJT in assessing SIs: as the participant knows the actual situation, it is very unlikely she will calculate an SI. Therefore, the first desideratum for a new paradigm is that the actual situation is not known to the participants. In the previous section we saw that this is also desirable for making the question-answer pair a natural part of the story, linking the focus-determining question to the alternative relevance-determining conditional. Therefore, in the new experimental paradigm stories of the type exemplified in (12) are used, in which the actual situation is not known to the participant.

In section 6.1 of the previous chapter I discussed a specific problem that arises because the TVJT asks for a judgment about (the truth of) a sentence, instead of assessing the interpretation more directly. The problem is that participants might have considered the answers in the non-focus conditions to be partly true, as they did provide the correct answer to the question, even though the SI made them false. The second desideratum for a new

experimental paradigm for testing SIs therefore is that it measures the SI more directly than through a truth value judgment.

SIs are inferences through which addressees rule out a certain possibility of what the actual world is like: the possibility which could also be described by the stronger scalar term. We can measure whether participants made these inferences by asking them whether they consider this state of affairs to be possible. For the story type in (12), repeated here, such a question is (13):

(12) new story type
Marieke told her mother that Katja went searching for marine animals on the beach yesterday, and that her grandfather had told her that if she would find at least two marine animals, she would get ten bucks.
The mother said: "Oh, and what did Katja find?"
Marieke answered: "She found a crab or a starfish."

(13) Do you think it is possible that Katja found both a crab and a starfish? yes / no

A 'no' answer to this question indicates that the participant ruled out the situation that could also have been described by the stronger scalar term, so she calculated the SI. This is a much more direct way of tapping into SIs than through a truth value judgment. As the participant is asked to judge the possibility of a world in which *A and B* is the case, I call this task the *Possible World Judgment Task* (PWJT). In the experiments presented below I used this paradigm to test the focus-sensitivity of SIs and tease apart the effect of focus and alternative relevance.

4. Experiment 4: Testing focus and alternative relevance with the PWJT

4.1 Setup and items

Experiment 4 was set up in the PWJT paradigm as described in the previous section. In line with the proposed manipulations in section 2.2, I included three conditions: [+alternative relevance, +focus], [−alternative relevance, −focus] and [−alternative relevance, +focus]. The last condition, which is labeled condition 1, is the crucial one for teasing apart focus and alternative relevance. Alternative relevance was manipulated with conditionals (in the way discussed in 2.2), and focus was manipulated with explicit questions, as in Experiments 1 and 2. Examples of the conditions are given in (14)-(16),

differences between the conditions are marked in boldface for easy reference (but this marking was absent in the actual experimental items).

(14) Condition 1: [-alternative relevance, +focus]
Marieke told her mother that Laura went searching for marine animals on the beach yesterday, and that her father had told her that **if she would find a crab or a mussel**, she would get to stay up late that night.
The mother said: "Oh, and **what did Laura find?**"
Marieke answered: "She found a crab or a mussel."

(15) Condition 2: [+alternative relevance, +focus]
Marieke told her mother that Laura went searching for marine animals on the beach yesterday, and that her father had told her that **if she would find at least two marine animals**, she would get to stay up late that night.
The mother said: "Oh, and **what did Laura find?**"
Marieke answered: "She found a crab or a mussel."

(16) Condition 3: [-alternative relevance, -focus]
Marieke told her mother that Laura **and Barbara** went searching for marine animals on the beach yesterday, and that their father had told them that **the one who would find a crab or a mussel**, would get to stay up late that night.
The mother said: "Oh, and **who of them found a crab or a mussel?**"
Marieke answered: "Laura found a crab or a mussel."

Alternative relevance was manipulated by the requirement in the conditional. In condition 1 and condition 3, the requirement was finding a crab or a mussel, so it was enough to find one of the two in order to get the reward. As the questioner asked the question to check whether the requirement was satisfied (see section 2.3 above), the stronger scalar alternative *Laura found a crab and a mussel* was not relevant to her. In condition 2 the requirement was finding at least two animals, so the stronger scalar alternative was very relevant to the questioner.[2]

Focus was manipulated with explicit questions in the same way as in Experiments 1 and 2. In the two [+focus] conditions (condition 1 and condition 2) the question of the mother is about the direct object, causing *A*

[2] Alternatives to *at least two* that were used were *more than one* and *several* (Dutch: 'meerdere').

or B (here: *a crab or a mussel*) to be the focus of the target sentence. In the non-focus condition (condition 3), the question was about the subject and as a result *or* was part of the background of the answer.

Every story was followed by three questions, exemplified in (17):

(17) Question 1: What would Laura be allowed to do if she would find a crab or a mussel? ...
Question 2: Do you think Marieke knows exactly what Laura found? yes / no
Question 3: Do you think it is possible that Laura found both a crab and a mussel? yes / no

The third question was the PWJT question about the possibility of the *A and B* situation, which was critical for our current purposes. The first question was an open comprehension question about the story. It was always about the second part of the first sentence, the part that contained the conditional. This question was included to ensure that participants read the part of the story in which alternative relevance was manipulated and not just the dialogue which contained the focus manipulation, which would bias the data in favor of the focus manipulation. As the conditional varied over conditions the comprehension question had different versions in the different conditions.

With the second question an additional issue was investigated, the issue of speaker expertise. However, in order to keep the discussion in this chapter from becoming too confusing, I will discuss the reasons for including this question type and the results of it in chapter 8 of this dissertation. I will consider whether this question could have interfered with our current goals in the discussion of this experiment (section 4.6).

A side-effect of leaving out the situation and asking explicitly whether something is possible, is that expectations based on plausibility might interfere. For instance, a participant might consider it highly unlikely that Laura found both a crab and a mussel in general, irrespective of the use of *or*. To control for this, I introduced a control condition which was identical to condition 3, with the only difference that the answer that was given was a term answer, i.e. the VP containing *or* was elided. See (18): [3]

[3] In Dutch it is not necessary to have *do*-support (e.g. the answer 'Laura did' in the Dutch version was 'Laura.').

(18) Condition 4: (control)
Marieke told her mother that Laura and Barbara went searching for marine animals on the beach yesterday, and that their father had told them that the one who would find a crab or a mussel, would get to stay up late that night.
The mother said: "Oh, and who of them found a crab or a mussel?"
Marieke answered: "Laura did."

If participants answer 'no' to the SI-question because they considered the *A and B* situation impossible or very implausible in general, this control condition should trigger many 'no' answers too. So by including this condition and comparing it to condition 3, we can straightforwardly see how the VP with *or* affects the possibility of the *and*-situation.

4.2 Predictions

I already discussed the predictions in section 2.2, but I will apply them here to the conditions presented above. Based on the TVJT-results, the claims about focus-sensitivity of SIs, and the claims about the necessity of alternative relevance, we expect the [+alternative relevance, +focus] condition 2 to trigger more SIs than the [-alternative relevance, -focus] condition 3. The crucial condition for teasing apart focus and alternative relevance is condition 1, which is [-alternative relevance, +focus]. If this condition patterns with condition 2 (as both are [+focus]) and triggers a higher SI-rate than condition 3, this suggests focus is the crucial property. If however condition 1 patterns with condition 3 (as both are [–alternative relevance]) and returns a lower SI-rate than condition 2, this suggests alternative relevance is what is important for SIs. If the SI-rate of condition 1 is in between those of condition 2 and condition 3, this indicates that both properties increase SIs. These predictions are summarized in (19).

(19) Predictions for Experiment 4:
(C1=C2) > C3: Focus is important for SIs
C2 > (C1=C3): Alternative relevance is important for SIs
C2 > C1 > C3: Both focus and alternative relevance are important for SIs

4.3 Design

12 story quadruples like the one in (14)-(16) and (18) were created (fill list in Appendix 4), and distributed over 4 lists, so every list contained 12 test items, 3 per condition. Every list also contained 4 fillers, which were similar to the test items, but contained a different question instead of the SI-question (e.g. of the type *Do you think it is possible Bart didn't catch a spider?*). The number of fillers was deliberately chosen to be relatively small, to keep the experiment short. Each list started with an item from a different condition, to avoid order effects. There were always at least two items of other conditions (or fillers) between two test items of the same condition.

4.4 Participants and procedure

50 participants, all adult native speakers of Dutch with no prior knowledge of the topic, were recruited via e-mail and filled out a web-based questionnaire on their own computers. The questionnaire ran on WWSTIM (Veenker (2000)). Most of the participants were students or had a university degree. Participants were instructed that they would read 16 stories, followed by three questions per story, of which the first one was an open question, and the other two were yes/no-questions starting with *Do you think*. They were asked to answer the open question by typing the answer in the textbox directly under it, and to click the 'yes' or 'no' buttons at the yes/no-questions. The items were presented one by one, with the story and all three questions remaining on the screen until the participant had answered them and pressed the 'next' button. The instructions also mentioned that there were no right or wrong answers, and that participants should rely on their intuition. There was no time limit, but most participants completed the experiment in less than 20 minutes, with an average of 15 minutes.

4.5 Results

The results of Experiment 4 are summarized in Table 1:

Table 1: SI-rates Experiment 4

Condition 1 [-alt +foc]	Condition 2 [+alt +foc]	Condition 3 [-alt -foc]	Condition 4 [control]
67%	72%	63%	27%

All SI-rates were significantly different from chance. Even though more SIs were calculated in condition 2 than condition 3, and the items analysis showed this was the case in 10 out of the 12 items (see Appendix 4), this difference was not significant over participants or over items (over subjects: Wilcoxon signed rank test $z = -1.46$, $p = 0.073$, over items: $z = -1.41$, $p = 0.079$). All p-values reported here are one-tailed, in line with the predictions in (19) above. So focus and alternative relevance together did not lead to a significantly higher SI-rate in condition 2 than condition 3. Not surprisingly, as the SI-rate of condition 1 is in between those of condition 2 and condition 3, condition 1 did not differ significantly from either of them (C1-C2: $z = -0.78$, $p = 0.219$, C1-C3: $z = -1.26$, $p = 0.105$).[4] Condition 3 did differ significantly from the control condition 4, both in the subjects ($z = -4.60$, $p < 0.001$) and the items analysis ($z = -2.98$, $p = 0.002$). This indicates the 'no' answers were not due to participants' expectations about plausibility of the *and*-situation, but must have been caused by the presence of the constituent with *or*.

4.6 Discussion

The effect of focus on SIs found in Experiments 1-3 was not replicated with the PWJT-paradigm. The SI-rates of the focus conditions (67% and 72%) were comparable to the SI-rates found in Experiments 1-3 (resp. 73%, 67% and 85%), but the non-focus condition 3 returned a higher SI-rate (63%) than in Experiments 1-3 (resp. 55%, 41% and 55%). This is surprising considering that the discrepancy between story and target sentence that was present in the TVJT was no longer present in this paradigm, so the 'no' answer cannot have been due to the infelicity of the weaker term. I will discuss two possible explanations for why the effect was not replicated in the General Discussion (section 7). Although the SI-rate of condition 1 was in between those of conditions 2 and 3, these differences were not significant, so it was impossible to draw any conclusions about the relative importance of focus and alternative relevance.

I decided to conduct a follow-up experiment to control for a number of possible sources of experimental noise in Experiment 4. The first possible source of experimental noise was some differences between the stories of the two conditions. While in conditions 1 and 2 one character was introduced (e.g. Laura), in condition 3 two characters were introduced (Laura and

[4] In the cases in this chapter in which I only report one analysis this is the subjects analysis. In these cases the analysis over items also did not return a significant result.

Barbara), in order to make the *who*-question possible. Also, another form of the conditional was used in condition 3 (e.g. *the one who would find*) than in conditions 1 and 2 (e.g. *if she would find*). These two differences between the stories of condition 3 on the one hand and conditions 1 and 2 on the other might have masked the predicted difference between them based on focus (more SIs in C1 and C2 than C3, see (19) above). Therefore, I decided to make the items of the follow-up experiment as similar as possible between the conditions, as I did in Experiment 2. I introduced two characters in all conditions and used the same form of the conditional in all conditions.

A second source of experimental noise I wanted to control for was the speaker expertise question that preceded the SI-question in the experiment. Although I do not see how this could have increased SI-rates (if anything I would expect a decrease, see chapter 8), it might have interfered in some way with a natural interpretation of the target sentence.

Thirdly, the conditional that was used in the [-alternative relevance] conditions might have decreased the plausibility that both A and B were the case. For instance in the examples above, the girls had no incentive to go on searching once they found one of the two required animals, as they were told beforehand that one would suffice. This might have led participants to consider it unlikely that they went on searching and found both. However, if this was the case participants should have also considered it unlikely in the control condition. Although the 27% 'no' answers in the control condition is much lower than the rate of 'no' answers in conditions 1 and 3, it could possibly have increased the rates in conditions 1 and 3 compared to condition 2 enough to mask the predicted effect of alternative relevance (more SIs in C2 than in C1 and C3, see (19) above). Therefore, this problem was also fixed in Experiment 5 by slightly changing the story.

5. Experiment 5: Controlling for three sources of experimental noise

5.1 Setup and items

The same four conditions were used as in Experiment 4. Example items of the conditions are given in (20-23). Again, differences between the conditions are given in boldface here, but not in the actual items.

(20) Condition 1: [–alternative relevance, +focus]
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **a crab or a mussel** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **what had Laura found** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(21) Condition 2: [+alternative relevance, +focus]
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **at least two marine animals** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **what had Laura found** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(22) Condition 3: [-alternative relevance, -focus]
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **a crab or a mussel** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **who of Laura and Barbara had found a crab or a mussel** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(23) Condition 4: (control)
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found a crab or a mussel on the beach, that person got to stay up late that night.
The friend said: "Oh, and who of Laura and Barbara had found a crab or a mussel on the beach?"
Marieke answered: "**Laura did.**"

Similar to Experiment 4, alternative relevance to the hearer was manipulated by the requirement in the conditional and focus was manipulated by an explicit question. Except for these manipulations, the conditions were identical. They all contained two characters that went out to do something

and they all contained the same form of the conditional. To prevent that the items were too boring, I used three types of conditionals: *if one of them... that person* (Dutch: 'als één van hen... diegene'), *the one who* (Dutch: 'degene die') and *if one* (generic) (Dutch: 'als je'). The form of the conditional was always kept constant between the different conditions of a story.

In the items exemplified in (20)-(23), the characters in the story did not know about the conditional until they returned, so this could not have influenced their behavior. Therefore, it is no longer implausible that they went on searching after they found one animal. As a result of this change, the tense of the question and the target sentence had to be changed from present perfect to past perfect. (Dutch: 'heeft' (*has*) was changed to 'had' (*had*)).

The stories were followed by two questions, as in (24):

(24) Question 1: What would the person who had found a crab or a mussel be allowed to do? ...
Question 2: Do you think it is possible that Laura had found both a crab and a mussel? yes / no

The first question was again a comprehension question, which was always about the second sentence. The second question was the SI-question. The speaker expertise question was left out in this experiment to avoid possible interference.

Similar to Experiment 4, Experiment 5 investigated an additional issue. Next to being a control experiment for Experiment 4, the experiment aimed at testing the relation between SIs and exhaustivity. Again, I will not go into this issue here, but I discuss the reason for testing this and the results in chapter 9. However, because of this additional issue half of the items was followed by (a comprehension question and) an SI-question like Question 2 in (24), and the other half was followed by (a comprehension question and) an exhaustivity question like (25):

(25) Do you think it is possible that Laura also found something else than a crab or a mussel? yes / no

I will discuss the results of the exhaustivity items in chapter 9, where I will also discuss possible carry-over from one type of item to the other. In this chapter I will ignore the exhaustivity items.

5.2 Design

Due to the fact that Experiment 5 also tested for exhaustivity, the number of SI-items per condition was reduced from 3 to 2. Of the 12 story quadruples of Experiment 4, 8 were re-used in Experiment 5 as SI-items. The remaining four and the fillers of Experiment 4 were turned into exhaustivity-items (see Appendix 5). The 8 story quadruples were distributed over 4 lists, so every list contained 8 SI-items (2 per condition). To maximize comparability to Experiment 4, the order of the stories per list was the same as in Experiment 4.

5.3 Participants and procedure

68 adult native speakers of Dutch participated in the experiment, which was conducted on computers in the lab. Most participants were undergraduate students of Utrecht University. None of them had any prior knowledge about the topic (e.g. none had followed an introductory course in logic or semantics). They were paid for their participation. The same software was used as in Experiment 4 (WWSTIM, Veenker (2000)). The procedure and instructions were the same as in Experiment 4, with the difference that now only two questions per item had to be answered. Participants were instructed to type the answer to the open comprehension question in the textbox and click on the 'yes' or 'no' button at the yes/no-question. It was pointed out to the participants that at the *Do you think*-question, they were supposed to follow their intuition. Again, the story and the questions remained on the screen until the participants had answered both questions and pressed the 'next' button. All participants completed the experiment in less than 20 minutes.

5.4 Results

The results of Experiment 5 are summarized in Table 2:

Table 2: SI-rates Experiment 5

Condition 1 [-alt +foc]	Condition 2 [+alt +foc]	Condition 3 [-alt -foc]	Condition 4 [control]
71%	76%	65%	18%

The pattern of the results is very similar to that of Experiment 4. However, here the difference between condition 2 and condition 3 was significant over

subjects and items (over subjects: Wilcoxon signed rank test z = -2.28, p = 0.012 , over items: z = -2.05, p = 0.021 (again all p-values are one-tailed, and the scores per item are given in Appendix 5)).[5] So focus and alternative relevance together had an effect on SI-calculation. However, condition 1 did not differ significantly from either condition 2 or condition 3 (C1-C2: over subjects: z = -1.36, p = 0.088 , over items: z = -1.47, p = 0.070. C1-C3: over subjects: z = -1.257, p = 0.104, over items: z = -0.95, p = 0.170). Again, the 'no' answers in the test conditions were not only due to plausibility: the SI-rate of condition 3 was significantly higher than the control condition 4 (over subjects: z = -5.55, p < 0.001, over items: z = -2.53, p = 0.006).

5.5 Discussion

After three possible interfering factors of Experiment 4 were controlled for (differences between stories in different conditions, the speaker expertise question, incentives of characters in the story), and when the experiment was conducted in a well-controlled environment (the lab), the experimental manipulations did result in a significant difference between the condition with both focus and alternative relevance and the condition with neither. This supports the view that the cumulative property of focus and alternative relevance increases SI-calculation. Again, the SI-rate of the [-alternative relevance, +focus] condition (C1) was in between the other two rates, but did not differ significantly from either of them, making it impossible to draw conclusions about the relation between the two. This raises the question whether the experiment was sensitive enough. I will return to this question in the General Discussion (section 7).

One of the reasons to switch to the PWJT was to avoid rejections based on the infelicity of using a weaker term, which might have led to an overestimation of the SI-rate in the TVJT experiments. However, the PWJT experiments actually returned an even higher rate of SIs, especially in the non-focus condition. Before I look into this in more detail in section 7, I present the SI-data of two experiments that were conducted as control experiments for the exhaustivity-data of Experiment 5 (see chapter 9). Experiment 6 was a paper-and-pencil version of Experiment 4, and Experiment 7 was a paper-and-pencil version of Experiment 5.

[5] The difference in the analysis over participants is still significant when we correct for multiple (4) comparisons. The Bonferroni corrected p-value is 0.046.

6. Experiments 6 and 7: Paper-and-Pencil control experiments

6.1. Experiment 6

6.1.1 Setup

Experiment 6 was set up as a control experiment for the comparison between SIs and exhaustivity in Experiment 5 (see chapter 9). Again, both SI and exhaustivity were tested. However, this time the experiment used a between-subjects design in order to control for the fact that Experiment 5 used different stories for SI-items and exh-items. The items and design of Experiment 4 were used for both versions of the experiment. One group of participants got the SI-version, containing only SI-questions, while another group got only exh-questions, with the same items. As a result of this setup, the SI-version of the experiment was a (paper-and-pencil) repetition of Experiment 4, but without the speaker expertise question. I will consider the setup and results of the SI-version here.

Participants were undergraduate students of Utrecht University. They filled out the questionnaire in class. All participants were native speakers of Dutch and had no prior knowledge of the topic. They were students of the Faculty of Humanities and all were taking the same (unrelated) course. 63 students filled out the SI-version of the experiment. The instructions were the same as in Experiment 4, with an additional instruction that participants were discouraged to look back at earlier items. This was done in order to increase similarity to the computer-based experiments, in which items were presented one by one on the screen.

6.1.2 Results

The results of Experiment 6 are summarized in Table 3:

Table 3: SI-rates Experiment 6

Condition 1 [-alt +foc]	Condition 2 [+alt +foc]	Condition 3 [-alt -foc]	Condition 4 [control]
68%	74%	67%	29%

The observed SI-rates are very similar to the results of Experiments 4 and 5. As in Experiment 5, the difference between condition 2 and condition 3 was significant, however only in the subjects analysis: $z = -1.69$, $p = 0.045$ (items analysis: $z = -1.37$, $p = 0.085$, see Appendix 4 for the rates per item).

Condition 1 again returned an SI-rate in between condition 2 and condition 3, but the differences were again not significant: (C1-C2: z = -1.27, p = 0.204, C1-C3: z = -0.28, p = 0.391). As in Experiment 5, if *or* is in the focus part of the sentence and the stronger alternative is relevant to the hearer, more SIs are calculated than when these two properties are absent, but we did not find evidence for which one of the two is more important.

6.2 Experiment 7

6.2.1 Setup

Experiment 7 was a paper-and-pencil version of Experiment 5, so again both SIs and exhaustivity were tested, but like Experiment 5 in a within-subjects design. While Experiment 6 used the items of Experiment 4, in Experiment 7 the items of Experiment 5 were used. The design was also adopted from Experiment 5.

34 undergraduate students of Utrecht University filled out the questionnaire after an exam. None of them had participated in any of the previous experiments or had any prior knowledge of the topic, and all of them were native speakers of Dutch.

6.2.2 Results

The results of Experiment 7 are summarized in Table 4:

Table 4: SI-rates Experiment 7

Condition 1 [+foc -alt]	Condition 2 [+foc +alt]	Condition 3 [-foc -alt]	Condition 4 [control]
65%	78%	68%	16%

Although the SI-rates of the different conditions were similar to the previous experiments, for the first time the SI-rate in condition 1 was lower than in condition 3, contrary to the prediction that focus increases SIs. The difference between condition 2 and condition 3, representing the cumulative effect of focus and alternative relevance, was not significant (over subjects: z = -1.32, p = 0.094, over items: z = -1.40, p = 0.081, all p-values one-tailed). However, the difference between condition 1 and condition 2 was significant over participants: z = -1.66, p = 0.049, and marginally significant over items: z = -1.54, p = 0.062 , showing a significant effect of alternative relevance.

7. General Discussion Experiments 4-7

In Table 5, the results of Experiments 4-7 are summarized:

Table 5: Summary SI-rates Experiments 4-7

	Condition 1 [-alt, +foc]	Condition 2 [+alt, +foc]	Condition 3 [-alt, -foc]	Condition 4 [control]
Exp 4	67%	72%	63%	27%
Exp 5	71%	76%	65%	18%
Exp 6	68%	74%	67%	29%
Exp 7	65%	78%	68%	16%

The significant differences between conditions 2 and 3 in Experiments 5 and 6 suggest that focus and alternative relevance together increase SI-calculation. However, the differences between condition 1 and the other test conditions is too small to draw any clear conclusions about the mutual relation of the two properties.

The fact that many differences were not significant raises the suspicion that the experiments were not sensitive enough to clearly reveal the effects. Perhaps the differences between the conditions were so subtle that the participants simply did not notice them. A strong argument against this explanation is that the exhaustivity-data, which will be discussed in chapter 9, did show clear differences between the conditions. As in Experiments 5 and 7 the same subjects answered the SI-items and the exh-items, we can conclude that the participants were aware of the differences between the conditions. Therefore, the explanation has to be specific to SIs. I will discuss two possible explanations below. As I find the second explanation to be the most plausible one, I discuss how we can avoid this problem and still avoid the problems of the TVJT in section 7.3.

7.1 An interfering Manner Implicature in the non-focus condition

In comparison with the TVJT-data, the SI-rate of the non-focus condition is unexpectedly high. One explanation for this comes from comparing condition 3 to the control condition 4. In condition 4 the SI-rate dropped dramatically. This is not surprising, as the target sentence in condition 4 did not contain *or*. However, the fact that in condition 3 the term answer (e.g. 'Laura did.') would have also been a possible answer, might have led to a Manner Implicature. A Manner Implicature is the inference from a marked way of saying something to a marked meaning, based on Grice's maxim of Manner (see chapter 1). The idea is that because there is a shorter way to

answer the question (the term answer), it is somehow marked to repeat the rest of the sentence, the part containing the disjunction. This might have triggered assumptions that either the speaker considered the direct object to be relevant (or questioned), or that by repeating the disjunction the speaker wanted to indicate that as far as she knows, Laura only found one of the two (the SI). These assumptions could have led to the high SI-rates in condition 3.

The question then arises why this effect was not observed in the TVJT-experiments. Notice that in the items of Experiment 2, the term answer and the sentential answer could not be distinguished, as indirect speech was used. For instance, both the term answer in (26) as the sentential answer in (27) can be described in indirect speech by (28):[6]

(26) mother: "Who found a crab or a starfish?"
Karin: "Julie did."

(27) mother: "Who found a crab or a starfish?"
Karin: "Julie found a crab or a starfish."

(28) Mother asked who had found a crab or a starfish.
Karin answered that Julie had found a crab or a starfish.

Therefore, the Manner Implicature did not arise in Experiment 2. In Experiment 1 however, the Manner Implicature could have also played a role (as direct speech was used), but there it was less salient as the speakers in that experiment were not part of the story. Moreover, contrary to Experiments 4-7, the TVJT-experiments did not contain a control condition in which the term answer was used, which might have highlighted the use of the sentential answer and facilitated the Manner Implicature. I will now consider an explanation that I find more compelling, which is based on the experimental paradigm itself.

7.2 Experimental question overrules critical manipulation

Another explanation for why no clear differences were observed in the PWJT experiments while there were in the TVJT experiments, is the task itself. It has been noted by several authors that a task in which the SI is

[6] The answer in (26) can also be described in indirect speech by *Karin answered that Julie did*. The point is that (28) is a *possible* description of (26) in indirect speech and therefore it cannot be derived from (28) which form of the answer was used, the term answer or the sentential answer.

explicitly asked about, is likely to return unrealistically high SI-rates. For instance Geurts & Pouscoulous (2009) reported SI-rates for an inference task, in which they explicitly asked whether the SI followed from the sentence, that were almost twice as high as in a TVJT with the same sentences. Geurts (in prep.) discusses the following example (29), with its SI in (30):

(29) Some of the goats have the flu.

(30) Not all of the goats have the flu.

He argues that once we ask ourselves (or participants for that matter) whether (29) implies (30), we bias our judgment because by asking this we make the question whether the stronger alternative holds relevant. He says:

> 'Obviously, to ask oneself whether or not (14a) [here (29), AZ] implies (14b) [here (30), AZ] is to suggest already that it *might* be implied, but more importantly, this question raises the issue whether or not all of the goats have the flu, in other words, it makes it *relevant* to establish whether this is the case.' (p. 94, his italics).

> '[...] when we consult our intuitions about (14a), the question is asked explicitly. Hence, our intuitions are not about how this sentence would be interpreted in general, but are confined to a limited range of contexts, that is, contexts in which it is relevant to establish if all goats have the flu.' (p. 94)

The point made by Geurts is that asking about the SI automatically makes the question whether the stronger alternative holds relevant. Frazier (2009) makes a related remark about the paraphrase selection task, in which participants have to choose between the reading of the sentence with and without SI. She says:

> 'Paraphrase selection may itself encourage readers or listeners to consider the various interpretations of the sentence even if they wouldn't have during a simple comprehension task without the paraphrase selection task.' (p. 326)

Frazier's point is that the experimental question can encourage participants to consider a certain interpretation, which they would have not considered spontaneously.

Something along the lines described by Geurts and Frazier might have been going on in Experiments 4-7 too. In the PWJT, the experimental question might have overruled the critical manipulations, which would explain why they were not effective. In line with what Geurts described above, by asking about the *A and B* situation the experimental question made the stronger scalar alternative relevant even in the [-alternative relevance] conditions. Similarly, as the experimental question was always about the direct object, the scalar term was always in the part of the sentence that was the information focus relative to the experimental question, even in the [-focus] condition. Therefore, due to the experimental question all three conditions became [+alternative relevance, +focus] conditions, overruling the critical manipulations.

The explanation that the experimental question caused the difference between the TVJT and the PWJT results becomes more clear if we compare the process the participant has to go through in the TVJT to that in the PWJT. In the TVJT, a participant reads the story in which *A and* B is the case. Then she reads the explicit question (the critical manipulation). Then she reads the target sentence of which she knows she will have to judge the truth value. At the moment the participant is interpreting the target sentence, the most recent indication of what is relevant is the explicit question, which is the critical manipulation of the experiment. This results in significantly different behavior between the two conditions. Even though there is a discrepancy between situation and target sentence in both conditions, the acceptability of this discrepancy is influenced by the explicit question.

However, in the PWJT the process is very different. The participant reads the story in which the actual situation is not given. She then reads the explicit question and the target sentence. Then she reads the experimental question about the possibility of the *A and B* situation. Now she has to go back to the story and check the target sentence with the question about the *A and B* situation in mind. Therefore, the most recent issue that is made relevant here is not the explicit question that was posed in the story (the critical manipulation), but the experimental question about the *A and B* situation. As this question was the same in all three conditions, it is not surprising it returned approximately the same rate of rejections in all three conditions. So actually we have been testing the question-answer pair in (31) in all three conditions:

(31) Q: Do you think it is possible that Laura had found both a crab and a mussel?
A: Laura had found a crab or a mussel on the beach.

In this light it is not surprising the rejection rate in all conditions was high.

Another argument in favor of this explanation is the difference between condition 3 and condition 4. In condition 4 the target sentence was the term answer, e.g. *Laura did*. This was not a possible answer to the experimental question, so participants had to consider the question in the story it was an answer to, resulting in a much lower rate of 'no'-answers than in condition 3. I conclude that it is very plausible that the experimental question in the PWJT overruled the critical manipulations and that this is why the effect was not replicated.

However, one could also argue that the results of the TVJT were not to be trusted. In 6.1 of the previous chapter I discussed an alternative explanation of the data due to the fact that the TVJT asks for a judgment. As the answer with SI in the non-focus conditions is arguably partly true, while this is not the case in the focus conditions, this could account for the effects which I attributed to a difference in SI-calculation. In the PWJT this difference between the conditions was absent, as it did not ask for a truth value judgment. This could explain why we did not find the same effect in the PWJT: there is no effect. Even though this goes against the significant differences we did find in Experiments 4-7, it is a viable hypothesis as the differences were small. To rule out this hypothesis we have to do additional experiments in an experimental paradigm that avoids both the problems of the TVJT as those of the PWJT. In the next section I will consider how this can be done.

7.3 How to avoid the problems of the TVJT and PWJT at the same time

Let us take stock. In the previous chapter I have argued that the TVJT was not very suitable to test SIs, because the actual situation is known to the participant. These are unnatural circumstances for an SI to arise and it causes the use of a weaker scalar item to be infelicitous. I tried to solve this problem with the PWJT, in which the actual situation was taken out and participants were asked about the possibility of the *A and B* situation. However, this experimental question might have become the issue relative to which the target sentence was interpreted, overruling the critical manipulations of the experiment. The experimental question caused the stronger alternative to be relevant and *A or B* to be the focus in all experimental conditions, blocking the effect observed in the TVJT-experiments. Therefore, we need an experimental paradigm in which the *A and B* situation can be left out altogether. We do not want the *A and B* situation to be given as in the TVJT, and we also do not want to explicitly ask about it as in the PWJT. A way to satisfy these requirements is to look at the on-line processing of sentences containing a scalar term. If there is a spontaneous and measurable on-line

reflex of SI-calculation during processing, we can leave out both the actual situation and the interfering experimental question, and still compare the conditions of interest. I will take this route and present two on-line experiments in chapter 6.

But before we consider the on-line experiments, we have to establish what the on-line reflex of SIs could look like. Fortunately, a number of authors have claimed to have found on-line effects of SI-calculation. In the next chapter I will consider some of these studies. These are relevant for the current enterprise because they provide the crucial predictions of our on-line experiments. Also, from these studies we can distill some crucial methodological issues that we have to take into account in our own on-line experiments. One set of experiments (the experiments by Katsos (2006) (and colleagues) are particularly relevant as they investigated the processing of scalar terms in SI-triggering versus SI-blocking contexts. Additionally, they found effects of information structure on the processing of scalar terms. As our hypothesis is also about the effect of information structure on SIs, the results of Katsos and colleagues are directly related to the current investigation.

Most of the on-line studies that I will consider in the next chapter were conducted to provide evidence in the two debates that have dominated the literature on SIs in the last decade: the globalist-localist debate and the defaultist-contextualist debate. In order to be able to understand the on-line studies I will also introduce these debates in the next chapter.

CHAPTER 5: PROCESSING STUDIES ON SCALAR IMPLICATURES AND TWO HEATED DEBATES

1. Introduction

At the end of the previous chapter I argued that we need to turn to on-line experiments to test the hypothesis that SIs are focus-sensitive while avoiding the problems of the two off-line paradigms used so far (the TVJT and the PWJT). If SI-calculation has an observable effect on processing, we can test the hypothesis without including the actual situation in the experimental items or asking an experimental question which might overrule the critical manipulations. In this chapter I will consider a number of on-line studies on SIs from the literature which indicate that SI-calculation has a reflex in real-time language processing. As experimental work on this topic has increased greatly in the last couple of years, I will not present an exhaustive list but pick out some experiments that are relevant for the current enterprise.

Most on-line experiments on SIs in the literature were done to find evidence in two heated debates in the world of SIs: the *globalist-localist* debate and the *defaultist-contextualist* debate. Both of these debates mainly focus on the process of SI-generation, not on the eventual interpretation of the sentence, and that is why on-line data are needed to distinguish between the predictions. The globalist-localist debate focuses on the issue whether SIs are computed after the semantic meaning of the whole sentence has been determined (globalist), or whether SIs are computed during compositional semantics (localist). I will discuss this debate in section 2. Most of the on-line experiments that were designed to address this question use the *Visual World Paradigm* for eye-tracking (Tanenhaus et. al 1995). I will present a number of examples in section 3.

The defaultist-contextualist debate pertains to the issue whether SIs are generated by default whenever a scalar term is encountered, possibly to be canceled later on (defaultist), or whether SIs are only calculated if they are licensed by the context (contextualist). In section 4 I will address this debate in more detail. On-line experiments that address this question mainly relied on measures of processing cost. The assumption of these studies is that SI-calculation (and possibly also SI-cancellation) brings along an observable processing cost. As the two views make different predictions about the conditions in which SIs are calculated (or canceled), measuring processing cost can determine which view is right.

These results on processing cost are interesting for the current enterprise, as they provide a prediction to test the hypothesis of focus-sensitivity of SIs on-line. In the literature, the processing costs of SI-calculation/cancellation were measured by looking at answering times and reading times. The

answering time studies are presented in section 5. In section 6 I will consider reading time studies that manipulated structural (sentence-level) factors and found a delay in the reading of scalar terms which is possibly due to SI-calculation.

One series of reading time studies is particularly relevant for the current enterprise, as it explicitly considered the effects of the wider context on the processing of sentences containing scalar terms. These are the experiments of Katsos (2006) (partly also reported in Breheny et al. (2006)). In section 7 I present these in detail as they will form the starting point of the experiments in the next chapter. One of the experiments of Katsos (2006), Experiment 4, is particularly interesting as it tested predictions on the presence or absence of SIs based on information structure. I will discuss this experiment separately in section 8. There I will also address the relation of the results to the hypothesis of focus-sensitivity of SIs of the current work, as the two seem to be conflicting. Finally, in section 9 I discuss how I will use Katsos' experiments as a starting point for the on-line experiments presented in the next chapter.

2. The globalist – localist debate

Traditionally, it is assumed pragmatic inferences take as input the semantic representation of the whole sentence. Reconsider the nonce derivation by which according to many theories SIs are derived, repeated here in (1).

(1)
i. The speaker used the scalar term *or*.
ii. The speaker could have uttered the same sentence with the scalar term *and* instead of *or*, which would have been stronger/more informative (because the sentence with *and* entails the sentence with *or*).
iii. The sentence with the stronger scalar term *and* would have also been relevant.
iv. The speaker is trying to be as informative as possible (she is obeying the Maxim of Quantity).
v. Apparently, the speaker does not have evidence for the sentence with *and*.
vi. The speaker is well informed.
vii. Therefore, it is likely that the speaker considers the sentence with *and* to be untrue.

This derivation cannot be carried out before the sentence meaning is computed, as it is based on reasoning about sentences that the speaker

considers true or false, and alternative sentences in which other scalar terms were used. The idea is that after compositional semantics has provided the truth-conditional meaning of the sentence, pragmatic reasoning as in (1) takes place, taking the wider linguistic context into account as well as the situation in which a sentence is uttered. By adding pragmatic inferences (like SIs) based on this wider context to the literal meaning, the pragmatically enriched meaning of the sentence is formed. This view is called the *global(ist)* view as it is about how the sentence gets interpreted in the global context of the utterance, and its followers are correspondingly called *globalists*.

The globalist view is often associated with what Geurts (in preparation) calls the 'Standard Recipe' for SI-calculation, which is described somewhat simplified in the nonce derivation in (1). This account however runs into problems when a sentence contains more than one scalar term. Consider for example (2), from Sauerland (2004):

(2) Kai had the broccoli or some of the peas last night.

Here, the scalar term *some* is embedded under another scalar term: *or*. Intuitively, this sentence has two SIs, corresponding to the two scalar terms, see (3) and (4):

(3) Kai did not have the broccoli and some of the peas last night.

(4) Kai did not have all of the peas last night.

However, it is unclear how the SI in (4) is derived by a nonce derivation like (1). As on the globalist view we can only compare the whole sentence meaning to its alternatives, it is impossible to do a comparison for the second disjunct only, which is what we need to derive (4). Comparison with the alternative sentence in which both scalar terms are replaced by their stronger alternatives, leads to the SI in (5), which is too weak.

(5) It is not the case that Kai had the broccoli and all of the peas last night.

Authors supporting the globalist view have made proposals on how to account for these examples (e.g. by adapting the Standard Recipe, see Sauerland 2004, discussed in chapter 8). However, examples like (5) have led some other authors to abandon the global view, and propose *localist* theories of SIs. Chierchia (2004) proposed a radically different view for SI-calculation (see chapter 2 section 3.2), in which comparison to alternatives

happens locally, during compositional semantics, and in which enriched meanings of parts of a sentence can combine by function application. This way, embedded implicatures of sentences like (5) can be accounted for.

There is fierce discussion between globalists and localists (see e.g. Geurts & Pouscoulous 2009). Although for some sentences the two views make different predictions about which SIs should arise, often these sentences are so complex (e.g. due to the presence of several logical operators) that they are hard to assess experimentally. However, the two views clearly make different predictions about the process of SI-calculation, especially the time-course: globalists argue the SI is calculated after the literal sentence meaning, while localists claim the two go in tandem. Therefore, researchers started conducting on-line experiments to test these predictions. I will present some of these in the next section.

3. Testing globalism vs. localism: Visual World Paradigm experiments

A growing number of experimental studies is using the Visual World Paradigm (Tanenhaus et al. 1995), in which a number of objects or pictures are placed in a scene, and participants' eye-movements towards the different objects are measured by eye-tracking, while they interpret a target sentence. This paradigm is very suitable to determine which interpretations are entertained or preferred by a participant during language comprehension. The experiments on SIs using this paradigm focused mainly on the globalist-localist discussion, trying to answer the following two questions:

1. Is there an initial semantic stage, in which the literal interpretation of the sentence is entertained?
2. Are SIs calculated during sentence comprehension (localist) or after (globalist)?

The first attempt (that I am aware of) to answer these questions using the Visual World Paradigm was a study by Storto & Tanenhaus (2004). Interestingly, they looked at the scalar term *or*. They compared participants' eye-movements in 3x3 scenes with pictures of objects, schematized in Table 1 and Table 2, while participants listened to the target sentence (6).

Table 1: schema of 'or-early' condition of Storto & Tanenhaus (2004)

locks	grapes	roller skates
horses	oranges	roller skates
locks	bananas	rabbits

Table 2: schema of 'or-late' condition of Storto & Tanenhaus (2004)

locks	grapes	snakes
horses	oranges	camels
locks	bananas	rabbits

(6) The grapes or the oranges are next to some locks. Please click on those locks.

The bottom row in both displays was not important for the manipulation. The crucial manipulation is that in Table 1 (called the 'or-early' condition), the cells to the right of the grapes and the oranges have the same objects in them: roller skates. Storto & Tanenhaus predicted that if participants calculated the SI immediately upon hearing *the grapes or the oranges*, they knew they should look for a property that did not hold for both the grapes and the oranges. Therefore, before having heard the target *locks*, they would already look more at the two cells on the left (which were different for the grapes and the oranges), than the two cells on the right (which were the same for the grapes and the oranges). This anticipation was not possible in Table 2 (the 'or-late' condition), where both the objects on the left and the right were different for the grapes and the oranges.

The results showed that participants indeed converged on the target (the locks) earlier in the or-early condition than in the or-late condition. Crucially, while in the or-late condition participants only converged on the target after the sentence was completed, they did so in the or-early condition during the time the word *locks* was played, so before the sentence was finished. Storto & Tanenhaus conclude:

> 'It appears that *or* is given an exclusive interpretation already before the sentence containing the disjunction has been processed in its entirety, which clearly undermines the "extreme" alternative to our experimental hypothesis that many authors seem to have attributed to Grice.' (p. 14).

However, Storto & Tanenhaus also acknowledged that the effect was not as local as in a control condition in which *and* was used and participants already focused on the shared object-type while hearing *are next to*.

Summing up, Storto & Tanenhaus' results go against the globalist view, but leave room for discussion.

More results were obtained by Visual World studies that focused on the scalar term *some*. For instance Huang & Snedeker (2006) presented participants with a 2x2 scene with pictures described in Table 3, while playing target sentence (7):

(7) Point to the girl that has some of the socks.

Table 3: schema of a Huang & Snedeker (2006) item

boy with 2 socks	girl with 2 socks
boy with nothing	girl with 3 soccer balls

Huang & Snedeker argued that if participants calculate the SI for *some* immediately after hearing it, then before the target word (*socks*) is interpreted and the temporal ambiguity is resolved, they will be more likely to look at the girl with the two socks than the girl with the three soccer balls, as the latter girl has *all* of the soccer balls in the whole scene, which is incompatible with the SI of *some*.

The results showed that participants were slower in finding the target than in a comparable condition with *all* (and an adapted scene), which according to Huang & Snedeker is indicative of an initial literal interpretation of *some* (on which both pictures on the right in Table 3 are possible targets), but participants' looks to the target were already above chance before phonological disambiguation, indicating the calculation of the SI took place locally. In a follow-up study with a longer ambiguous period Huang & Snedeker (2007) found more evidence for this. So Huang & Snedeker's results support the global view in the sense that there seems to be an initial literal stage, but support the local view as the SI is calculated during sentence comprehension. However, Grodner, Klein, Carbari and Tanenhaus (2007) found conflicting results as they found no sign of an early literal interpretation. Summing up, the results of the Visual World Paradigm experiments have not yet led to the settling of the globalist-localist debate, but have provided some important data.

Although the Visual World paradigm is very suitable for testing the interpretation of sentences in isolation, it is less suitable for the current enterprise, in which the wider linguistic context is under investigation. In order to test the focus-sensitive accounts of SIs we have to manipulate that certain things are old or new information, which would certainly bias a comparison of gaze direction preferences between these conditions. Also, it is hard to come up with a good visual display for the type of sentences with *or* I have been investigating so far. Therefore, I decided to consider

processing cost differences caused by presence or absence of the SI. A number of experiments from the literature that have shown interesting results on the processing costs of SIs are discussed in sections 5-8 below. Most of these experiments were conducted to provide evidence for either side of another important debate in the literature on SIs: the defaultist-contextualist debate, which will be discussed in the next section.

4. The defaultist – contextualist debate

The defaultist-contextualist debate stems from Grice's (1967) distinction between Particularized Conversational Implicatures (PCIs) and Generalized Conversational Implicatures (GCIs). PCIs are *particularized* in the sense that they rely on the particular context and situation. They come about by a nonce derivation like (1) above. However, some conversational implicatures are so frequent that one could say they do not rely on some particular context or situation. Grice called these Generalized Conversational Implicatures, as these arise in general, or 'in absence of special circumstances':

> 'Sometimes one can say that the use of a certain form of words in an utterance would normally (in the absence of special circumstances) carry such-and-such an implicature or type of implicature.' (Grice 1967/1989: p. 37)

Horn (1972) argued the SI of *some* and *or* belong to this kind of conversational implicature, as this quote illustrates:

> '[...] in normal contexts, all things being equal, existentials are upper bounded by implicature, and disjunctions are exclusive by the corresponding implicature.' (p. 98)

As the generation of this kind of implicature does not need contextual support, it is claimed they are generated *by default* upon encountering for instance *or* or *some*. This view has been defended most famously by Levinson (2000).

However, default generation does not mean these SIs always stay, they can be implicitly and explicitly cancelled. Implicit cancellation happens when the SI would lead to a contradiction with the previous context or, to repeat Levinson's quote from chapter 1 'it is clear from the context of utterance that such an inference could not have been intended as part of the utterance's full communicative import' (Levinson 1983 p. 115-116). In that sense, the defaultist view is still context-sensitive: generation is default, but

whether or not an SI is present 'at the end of the day' depends on the context. In this sense, Levinson's model is a two-step model: the generation of SIs is default, but an additional Gricean layer determines whether they stay or get canceled (see Katsos 2006 for further discussion of the two-step nature of this model).

Proponents of the opposing view, called *contextualists*, claim SIs are not GCIs, but are in principle no different from other PCIs. They are only generated if they are licensed by the context. On this view, the generation of SIs is not default, but comes about by a nonce derivation like (1) above. This hypothesis is put forward most strongly by scholars working in the framework of Relevance Theory (Sperber & Wilson 1986, Carston 1998), although they replace Grice's terms with notions of their own (see chapter 1).

Similar to the globalist-localist debate, it is impossible to distinguish between the defaultist view and the contextualist view based on off-line data. If an SI is absent in a certain context, defaultists argue it was implicitly cancelled and contextualists argue it was not generated in the first place. The only type of context that would actually distinguish between the two views would be a totally neutral context. Defaultists predict SIs will be generated in these contexts, as no contextual support is needed, while contextualists predict the lack of contextual support will cause the SIs not to be generated. However, I do not think such neutral contexts actually exist. Even a sentence in isolation triggers many assumptions based on its lexical content and world knowledge. Also, information structure raises certain expectations of what came before (see section 5.3 of chapter 3 and the discussion of Experiment 4 of Katsos (2006) in section 8 below).

Fortunately, the defaultist view and the contextualist view make different predictions about the processing cost of SIs. On the defaultist view, generation of SIs is a default process, and is therefore cheap. Levinson (2000) for instance argues GCIs are a way of compensating the slow speaking rate of humans. He says: 'inference is cheap, articulation expensive, and thus the design requirements are for a system that maximizes inference' (p. 29). The contextualists, on the other hand, argue SIs are costly inferences, which will only be made if they are worth the effort. A number of researchers have tested whether SIs are costly or cheap by comparing answering times of answers reflecting SIs with answers reflecting literal interpretations. I will consider some of these in the next section.

5. Testing defaultism vs. contextualism with answering time experiments

As I discussed in the previous section, the defaultists and the contextualists make different predictions about the processing cost of SIs. The defaultists claim SIs are cheap, and cancellation might be effortful, while the contextualists claim SIs are costly. One way to test these predictions is considering the time it takes for participants to come up with a truth value judgment on sentences that potentially trigger an SI. Noveck (2001) tested (French equivalents of) generic sentences in which the scalar term *some* (French 'certains') was used, where the stronger *all* ('tous') would have also been appropriate, e.g. (8).

(8) Some elephants have trunks.

In 59% of the cases, the sentence was considered 'false', indicating the participant calculated the SI of *some*. Noveck & Posada (2003) and Bott & Noveck (2004) compared the time it took for participants to give a 'false' answer to sentences of this type to the time that was taken for a 'true' answer, in which the SI was apparently not calculated (or consequently cancelled). The contextualist view predicts a 'false' answer to take longer, due to the processing cost of the SI, while the defaultist view predicts a 'true' answer to take longer, due to the processing cost of canceling the SI. In both studies (Noveck & Posada (2003) and Experiment 3 of Bott & Noveck (2004)), participants took longer to answer 'false', in line with the predictions of the contextualist view.

De Neys & Schaeken (2007) support this result with an experiment in which participants had to carry out another task (memorizing dot patterns) while judging Noveck-sentences such as (8) above. When the load of the distractor task was increased, SI-answers went down from 79% to 73%. Also, the speed of the SI-answers went down as the load went up, while the speed of the non SI-answers stayed the same. These results also support the claim that the SI is costly, in line with the contextualist view.

A problem with the answering time experiments is that the measure is quite indirect. The time that is measured is not just the time taken to interpret the sentence (with or without SI), but also includes the whole decision-making process.[1] Therefore, one could argue it is not surprising that SI-

[1] A related issue is that SI-answers typically involve a 'false' response which might take longer than a 'true' response in general. Bott & Noveck (2004) acknowledge this possibility and conduct a control experiment for another experiment they present (which involved explicit instruction to interpret with or without SI), but not for Experiment 3 which is discussed here. However, this is just one of the factors of the decision-making process that might interfere with the processes of interest.

answers take longer. Once a participant takes the SI into account, there is a conflict between the two possible interpretations of the sentence (with and without SI): the meaning with SI is false, while the meaning without SI is true. This could very well slow down the decision-making process, compared to when the SI is not considered at all. So the difference between the two answering times might be due to something else than SI-generation.

Interestingly, Feeney et al. (2004) found that 'true' answers to Noveck-sentences (so answers without SI) also take longer than 'true' answers to sentences which are true irrespective of SI, like *Some birds live in cages*. This indicates that even in cases where participants eventually answered 'true', they might have entertained the SI-reading. Also, in Feeney et al.'s experiment, the effect found by Noveck and colleagues was not replicated.

More recently, Degen et al. (2009) found no delay of an SI-response. They measured answering times for sentences like (9) in a situation in which the participant got all of the gumballs.

(9) You got summa the gumballs.

They found that SI-answers ('no') were actually faster than non SI-answers ('yes'), which is in line with the defaultist predictions.

We can conclude that although some reading time experiments provided indications that SIs have a processing cost, which supports the contextualist view, some studies show different results, making it hard to draw conclusions. As I pointed out, the biggest problem of the answering time experiments is that the time measured is not just the time needed to interpret the sentence, but includes other processes as well, which might blur the view on the processing cost of SIs.

There is a variation on the answering time studies, which is to set the time given to participants to answer and vary it between conditions. Bott & Noveck (2002) (Experiment 4) found that for *some*, SIs went down (from 44% to 28%) when the allowed answering time went down, again suggesting SIs are costly. Chevallier et al. (2008) found very interesting results with the inverse paradigm for *or*. Participants that were made to wait for three seconds before answering calculated more SIs than participants who got to answer right away (48% vs. 25%). Chevallier et al. concluded that letting participants wait motivates them to enrich the meaning of the sentence. The problem of these experiments is that the time constraints might lead to a change in answering strategy which has nothing to do with the inference under study. In the Bott & Noveck experiment, participants under time pressure might guess that differences between *some* and *all* are of lesser importance and only pay attention to the lexical match between e.g. *elephants* and *trunks*, while participants in the slow condition might make

other assumptions. In the study of Chevallier et al., participants who got to answer right away could have skipped reading (or have not paid attention to) the second disjunct and answer 'true' if the first one was already the case.[2] For the participants that had to wait, this strategy was impossible. Also, the chances of participants noticing that both disjuncts were the case was increased in the delay condition, which could by itself explain the effect. Of course this is only speculation, but it illustrates how vulnerable this paradigm is to interfering factors due to the fact that an aspect of the task itself is changed between conditions.

A more direct measure of processing cost are reading times. I will now consider some experiments which measured reading times of scalar sentences in order to provide evidence for either side of the defaultist-contextualist debate.

6. Testing defaultism vs. contextualism with reading time experiments

Another point where the predictions of defaultist and contextualist theory diverge is in which type of context the interpretation of a sentence with a scalar term leads to higher processing load: in contexts that trigger SIs or contexts that block or cancel SIs. In the literature the first type of context, an SI-inducing context, is often called an *Upper Bound (UB)* context, as the upper bound of the meaning of the scalar term is important. The SI-blocking/canceling context is correspondingly called a *Lower Bound (LB)* context, as only the lower bound of the meaning of the scalar term is relevant there. The defaultists predict that in a UB context, the SI is calculated (by default), and in an LB context, the SI is calculated (by default) and consequently cancelled. Therefore, interpreting a sentence with a scalar term should lead to a higher processing load in an LB context than in a UB context, due to the additional process of canceling the SI. Contrary to this, the contextualists claim generation of the SI is costly, so the UB context, in which the SI is generated, should lead to a higher processing load than the LB context, in which according to them no SI is calculated to begin with. These predictions are summarized in (10) and (11):

(10) defaultists: UB: cheap generation
LB: generation + (possibly costly) cancellation
$\Rightarrow$ LB higher processing load than UB

[2] The fact that Chevallier et al. used existential sentences (*There is an A or a B*) which were obviously not exhaustive (see chapter 3 footnote 13) could have added to this strategy. This could also explain why the SI-rate, especially in the normal condition, was exceptionally low (25%).

(11) contextualists: UB: costly generation
LB: no generation
⇒ UB higher processing load than LB

These predictions have been tested in a number of experiments that compared the processing load of a scalar term in UB and LB contexts through reading times. In this section I will discuss a number of studies that have tested these predictions by manipulating structural properties of sentences, such as upward vs. downward entailment, to create UB and LB-environments. In the next section I will turn to the experiments that are most relevant for the current enterprise: reading time experiments in which the wider context was manipulated instead of structural (sentence-level) factors.

As is well-known from the literature, downward entailing (DE) environments block or flip SIs as the entailment relation between the items on the scale is reversed (see chapter 1). A number of studies have therefore compared reading times of segments containing a scalar term in a DE-sentence with reading times of the same region in an UE-sentence. Bezuidenhout, Morris and Wildmann (2009) compared reading times of regions containing *or* and *some* in the antecedents of conditionals and the restrictors of universals (which are both DE environments) with reading times of similar regions in the consequents of conditionals and the nuclear scope of universals (which are both UE). For example, they compared reading times of the segments given here in italics in (12) and (13), which contain the scalar term *some*: [1]

(12) DE: Every cook *who ground some of the peppers*, prepared a hot sauce.

(13) UE: Every cook who prepared a hot sauce, *ground some of the peppers*.

They found that the critical segment was read slower in UE environments, in line with the contextualist predictions.

One of the things that makes experiments like these hard to interpret, is that the UE-DE difference unavoidably introduces differences between the target sentences that the critical region is part of, which by itself might be responsible for the observed effect. Bezuidenhout et al. claim to have controlled for this bias by comparing the reading time difference between the scalar regions in the different positions to the reading time difference between the regions that did not contain a scalar term. For example, they compared the reading time difference between the scalar regions of (12) and

(13) to the reading time difference between the regions in the same position that did not contain the scalar term, given in italics in (14) and (15).[3]

(14) Every cook *who prepared a hot sauce*, ground some of the peppers.

(15) Every cook who ground some of the peppers, *prepared a hot sauce*.

Bezuidenhout et al. found that the reading time difference between (12) and (13) was bigger than between (14) and (15), and concluded this was due to the SI in (13). However, it is still the case that *the same lexical material* is only compared between two positions. So if for some unrelated reason *ground some of the peppers* is harder to process in the nuclear scope than in the restrictor of a universal quantifier, this effect will only show up in the comparison of (12) and (13), and comparing to the difference between (14) and (15) will not control for this. Better control conditions would have been equivalents of (12) and (13) with *some* replaced by *all*. This would allow for a comparison of the reading time difference between *some* and *all* in the restrictor versus the difference between *some* and *all* in the nuclear scope. That would be a comparison of the same lexical material (except for *some/all*) in the same structural environment.

Many UE vs. DE reading time studies run into problems like these. Panizza, Chierchia & Clifton (2009) compared reading times on the numeral (the alleged SI-trigger) in the Italian versions of (16) and (17):

(16) UE: John has two cars in the garage and he parks a motorcycle in the courtyard.

(17) DE: If John has two cars in the garage, he will park a motorcycle in the courtyard.

They found that the numeral was read slower in UE contexts, which they claim is due to the SI-computation. However, there is no a priori reason to

[3] Strangely, Bezuidenhout et al. chose bigger regions than just the NP containing *some* or *or*, which resulted in a difference in the number of words between the conditions.

assume the critical region would be read equally fast in a normal episodic sentence as in the antecedent of a conditional.[4]

The UE vs. DE experiments are not directly relevant for the current enterprise, as they compare sentences in isolation, where a structural (sentence-level) property is varied, while we are interested in contextual properties. However, they do show indications that the processing cost of an SI can indeed be measured on-line. One set of reading time experiments from the literature did manipulate context instead of structural properties, contrasting reading times of segments containing scalar terms in Upper Bound (UB) contexts versus Lower Bound (LB) contexts. These are the experiments by Katsos (2006). I will turn to these now.

7. Reading time experiments on Upper Bound vs. Lower Bound contexts: Katsos (2006)

The most elaborate set of reading time experiments in which UB (SI-triggering) and LB (SI-blocking) contexts were manipulated was carried out by Napoleon Katsos and colleagues. A number of these experiments were published in Breheny, Katsos and Williams (2006), but I will use the numbering of experiments from Katsos (2006). In this section I will consider Experiments 1-3 and Experiment 5 in detail in 7.1-7.4. These experiments suggest that the extra processing cost of SIs can also be observed in reading times if we contrast UB and LB contexts. Experiment 4 is especially relevant for the current enterprise as it tested predictions about SI-calculation based on information structure. Therefore, I discuss this experiment and its relation to the hypothesis of the current work separately in section 8. In section 9 I

[4] Panizza et al. claim the effect cannot be explained as 'a general influence of a specific grammatical construction since we tested two different environments (conditionals and [universal] quantifiers)' (p. 30). However, earlier in their paper they spell out the truth conditions of the conditional in (17) as in (i):

(i) In any situation in which John has two cars...

This already gives away the semantic similarity between the two constructions they tested (conditionals and universal quantifiers): in both cases the numeral occurred in the restrictor of some universal quantification. However, even if two genuinely different constructions were tested, this does not mean the effect cannot be due to the differences between those constructions and the UE environment. Again, there is no a priori reason to assume that the critical region will be read equally fast in UE episodic sentences as they are read *on average* in restrictors of universal quantifiers and antecedents of conditionals.

consider how I will use the experiments of Katsos and colleagues as a starting point for the experiments in the next chapter.

7.1 Katsos (2006) Experiment 1: Greek *or* in UB and LB contexts

Experiment 1 of Katsos (2006) (also Experiment 1 in Breheny et. al (2006)) is frequently referred to in the literature. This experiment was originally set up as a redo of an experiment by Bezuidenhout & Cutting (2002), who did a line-by-line Self Paced Reading (SPR) study with sentences containing a conversational implicature trigger. They compared reading times in what they called 'minimal' (LB) contexts to 'enriched' (UB) contexts. However, they used 6 types of implicature triggers, only one of which might be considered an SI-trigger: numerals. Bezuidenhout & Cutting found that the same sentence containing a conversational implicature trigger was read slower in LB contexts than in UB contexts, which is in line with the defaultist predictions.

Katsos and colleagues set up a similar experiment in Greek using only one implicature trigger, the (Greek equivalent of the) scalar term *or*. They compared reading times on segments containing *or* in a segment-by-segment fixed-window SPR paradigm in UB contexts and LB contexts. Example items are given in (18) and (19), where slashes indicate segment breaks, and the critical region is given here in italics.

(18) UB context
John asked / why Mary was entitled to a bonus, / and he wasn't. / Theo said / that Mary satisfied one of the prerequisites, / only he couldn't exactly remember which, / she had / *a masters or three years of related experience*. /

(19) LB context
Mary had a look at the advertisement for the job. / She wanted to make sure / that she had the formal qualifications / to apply. / The advertisement read: / you should have / *a masters or three years of related experience.* /

They found that, in line with the contextualist view, the critical segment was read slower in the UB context than in the LB context. This finding has often been cited as evidence for the processing cost of SIs, and for the contextualist view.

There is however a number of serious problems with this experiment. Firstly, in the UB context in (18), the context already rules out an inclusive

interpretation of *or*: from *she satisfied one of the prerequisites but he couldn't remember which*, we can already conclude she does not have both a masters and three years of experience. Therefore, there was no need to calculate the SI in (18). This was the case in 9 of the 12 UB contexts that were used in the experiment. In some items the exclusivity of *or* was given by explicit material, see the critical passages in (20)-(24), and in some items it was given by world knowledge (e.g. that someone cannot be born in two places), see (25)-(27):

(20) He had a hard time deciding who to talk to first, his brother or a friend of Mary.

(21) You can have the plat du jour for free. Today, customers could have for free meat or fish.

(22) John believes that one of the two is true: Peter must be English or must have lived in England.

(23) She is thinking what the most beneficial exercise would be. She finds it difficult to decide between going to the gym or going jogging.

(24) He stayed with a relative of his whom he hadn't seen for many years, his uncle or his cousin.

(25) Paul's mother tongue is German or French.

(26) He was wondering which language to write it in, Dutch or English.

(27) Theo was born and raised on an island of the Sporades, Skopelos or Skiathos.

So in the majority of the UB items, there was no need to calculate the SI. This makes the claim that the longer reading times were due to an SI less plausible.

The LB contexts were also problematic. As exemplified in (19), almost all (11 out of 12) target sentences in the LB contexts contained a modal verb. In 5 items the target sentence contained *should* or *must*, in 5 items it contained *advise* or *suggest* and in one item it contained *can*. These modals all create *any*-licensing environments, which as Chierchia (2004) points out, can block SIs. Therefore, it is questionable whether the intuitive absence of the SI in the LB contexts is due to the contextual manipulation, or due to

structural properties of the target sentence. This again raises the problem I pointed out in the previous subsection: we cannot assume identical segments are read at the same speed in different structural environments.

This touches on a more general problem of the experiment: the contexts and the target sentences were completely different in the two conditions, which could have introduced many differences between the conditions that were not controlled for. Also, it was unclear which properties of the context were manipulated. It seems they were generated based on general intuitions about their SI-triggering or blocking status. A final problem of the experiment was methodological: the critical region varied greatly in length (varying from 3 to 9 words). Moreover, it was sentence-final, bringing along unwanted sentence wrap-up effects. Summing up, although this experiment is much cited, it is questionable whether it actually measured what is was set up to measure. Fortunately, in the following experiment many of these problems were fixed.

7.2 Katsos (2006) Experiment 2: English *or* after different question-types

Experiment 2 was conducted in English and compared reading times of segments with *or* in three conditions: in LB vs. UB contexts, and in a DE-environment. Example items are given in (28)-(30), with the critical region given in italics.

(28) UB context
The director / asked his consultant: / Who is representing our company / at the court hearing? / His consultant replied: / *Turner or Morris* / from the Legal Department. /

(29) LB context
The director / asked his consultant: / Who is available to represent our company / at the court hearing? / His consultant replied: / *Turner or Morris* / from the Legal Department. /

(30) DE-environment (in UB context)
The director / asked his consultant: / Who is representing our company / at the court hearing? / His consultant replied: / I believe that if / *Turner or Morris* / from the Legal Department / do so, / we need not worry too much. /

In this experiment, whether the context was UB or LB was manipulated by the question that the target sentence was an answer to, and the target sentence was identical in both conditions. The DE condition was the UB question-context with the target segment embedded in the antecedent of a conditional, which is a DE-environment. The results again showed that the region containing the scalar term was read slower in the UB contexts than in the LB contexts. Also, the scalar region was read equally fast in the DE-condition as the LB context. This again confirms the predictions of the contextualists: SI calculation is costly and therefore UB contexts, in which the SI is calculated, cause a higher processing load than when the SI is blocked by the context (LB) or by structural factors (DE).

Many problems of Experiment 1 are solved in this experiment: the UB and LB contexts are minimally different and the target sentence is identical. Exclusivity is not already given in the context in the UB condition, as in (20)-(27) above. Also, the critical region is no longer sentence final, and as only names were used, there is no difference between items in the number of words in the critical region anymore. However, one problem from Experiment 1 remains. In the LB condition, the target segment is interpreted in an *any*-licensing environment. Even though the overt part of the target sentence is identical in the UB and LB contexts, the contrast between (31) and (32) shows that the part that is elided (the VP) still affects the *any*-licensing properties of the term answer.

(31) Q: Who is representing our company at the court hearing?
A: *Anybody.

(32) Q: Who is available to represent our company at the court hearing?
A: Anybody.

Therefore, it seems that in the LB context the SI-blocking is again triggered by a structural property rather than a contextual one. In this light, it is not surprising that the average reading time of the LB condition was very similar to that of the DE condition. Again, we run into the problem whether we can assume the critical region is read equally fast in different structural environments. As in Experiment 1, there is no check whether the slowdown in the UB contexts is due to the SI. Including control conditions for the three experimental conditions with *and* instead of *or* could have provided this check. As *and* is the top item of the scale, it triggers no SIs and would therefore be a good baseline condition. I will use this strategy in the experiments in the next chapter.

Finally, a weakness of both Experiment 1 and Experiment 2 of Katsos (2006) is that the eventual interpretation of the sentence is not assessed, i.e.

there is no way to tell whether the SI was actually calculated more often in the UB than in the LB contexts. The items of Experiment 2 were checked off-line in a paraphrase selection task on a seven point scale, but it is not obvious that participants will exhibit the same behavior in an off-line paraphrase task as during reading in an on-line task. It would be nice to assess the interpretation of the target sentence in the on-line experiments too. In Experiment 3, Katsos and colleagues found a smart way to do this. In the second experiment presented in the next chapter will adopt a variation of this.

7.3 Katsos (2006) Experiment 3: English *some* and facilitation on *the rest*

In Experiment 3, Katsos and colleagues switched to the scalar term *some* in English. In another study on *some*, Bezuidenhout & Morris (2004) compared reading times on a region in the next sentence to determine how the scalar term in the first sentence was interpreted. They compared sentences like (33) and (34):

(33) Some books had color pictures. In fact all of them did, which is why the teachers liked them.

(34) The books had color pictures. In fact all of them did, which is why the teachers liked them.

Bezuidenhout & Morris predicted that if the SI of *some* is calculated in (33), we should find a delay somewhere on *In fact all of them did* in the second sentence, as the SI is cancelled there. However, their results were hard to interpret as it was not completely clear where exactly in the region the cancellation took place.

Katsos and colleagues also used the strategy of measuring reading times of a region in the next sentence. They looked at the region *the rest*, as in (35) and (36):

(35) <u>UB context</u>
The young candidate's advisor / asked the campaign manager: / How is our candidate / doing in the polls? / The campaign manager replied: / He has managed / to overtake / *some of his opponents* / that have little funding. / *The rest* / of his opponents / are too far ahead of him.

(36) LB context
The young candidate's advisor / asked the campaign manager: / What must our candidate do / to get out / of the last position / in the polls? / The campaign manager replied: / He has / to overtake / *some of his opponents* / that have little funding. / *The rest* / of his opponents / are too far ahead of him.

The idea is that if the SI is calculated for *some of his opponents*, it is interpreted as *some but not all of his opponents*, which introduces a complement set of opponents (the opponents that were not overtaken / do not have to be overtaken), which is a salient antecedent for *the rest*. If however *some of his opponents* is interpreted without SI, as *some and possibly all of the opponents*, the set of opponents that were not overtaken / do not have to be overtaken has to be accommodated at *the rest*. Therefore, in the UB context reading of *the rest* should be facilitated compared to the LB context. Crucially, Katsos measured reading times on the scalar region itself (e.g. *some of his opponents*) as well as on *the rest*. This way it could be determined whether slower reading of the scalar region corresponded with faster reading of *the rest*, which would support the claim that the delay on the scalar region is due to the SI.

The results showed that the scalar region was read slower in the UB contexts than the LB contexts, while *the rest* was read faster in the UB contexts than the LB contexts. The faster reading of *the rest* indicates that the SI was indeed calculated more readily in the UB context, and the delay on the scalar term in this condition indicates that the SI-calculation was costly, supporting the contextualist view.

Although the setup of this experiment is very elegant, it still has some of the problems we saw in Experiments 1 and 2. As in Experiment 1, the contexts and the sentences containing the scalar term were quite different between conditions. Again, it was not clear which contextual properties were manipulated between the conditions. Also, all 18 LB contexts that were used contained a deontic modal verb (*has to / should*) in the sentence containing the scalar term, so again the scalar term was in an *any*-licensing environment in the LB condition. In Experiment 5, which will be discussed below, this problem was fixed.

7.4 Katsos (2006) Experiment 5: Greek *some* and facilitation on *the rest*

Experiment 5 (Experiment 3 of Breheny et al. (2006)), was conducted in Greek. Again, reading times of the region containing *some* as well as of the

the rest region in the following sentence were measured. Example items are given in (37) and (38).

(37) UB context
Mary / asked John / whether he intended to host / all his relatives / in his tiny apartment. / John replied / that he intended to host / *some of his relatives*. / *The rest* / would stay / in a nearby hotel.

(38) LB context
Mary was surprised / to see John / cleaning his apartment / and she asked / the reason why. / John told her / that he intended to host / *some of his relatives*. / *The rest* / would stay / in a nearby hotel.

Katsos also included a baseline condition, which was identical to (37) but contained *only some* instead of *some*. The results showed that again the scalar region was read slower in the UB contexts than in the LB contexts, and the *the rest* region was read faster in the UB contexts. In fact, it was read just as fast in the UB condition as in the baseline condition with *only some*, in which the upper bound of *some* was made explicit. These results again support the contextualist view that SIs are costly inferences which are not generated in LB contexts.

In this experiment some problems of Experiment 3 were corrected. The sentence containing the scalar term was almost identical in both conditions, and the LB context no longer contained a modal verb. However, the contexts themselves were still rather different between conditions, and it was still unclear which properties were manipulated. A new problem of Experiment 5 was that in all 18 UB contexts, exemplified in (37), the stronger scalar term *all* was present in the story. As a result, *some* in the target sentence gets a contrastive reading. This by itself could be responsible for the longer reading times. The explicit contrast with *all* makes *some* a very salient part of John's reply, which is not the case in the LB context. Also, one could wonder whether it is still needed to calculate the SI, or whether the negation of *all* is given by the fact that John corrects it in his answer by replacing it by *some*.

This problem carries over to the reading times on *the rest*. Due to the explicit contrast between *all* and *some* in the UB context (37), the complement set of relatives that were not intended by John to be hosted already becomes salient. It is questionable whether the SI is responsible for the facilitation of *the rest* or whether this is caused by the explicit contrast.

Finally, now that the LB contexts no longer contain modal verbs, they do not strike me as SI-blocking anymore. Unfortunately, Katsos did not report any off-line check of these items. For instance, I would be very tempted to conclude from the second sentence of (38) that John did not

intend to host all of his relatives. Therefore, it seems this experiment actually compared two types of UB contexts: one with an explicit contrast with *all*, and one without.

The final experiment of Katsos (2006) that I will discuss (Experiment 4) is directly related to the hypothesis of this work, as it checks predictions about how information structure leads to triggering or blocking of SIs. Therefore, I discuss it in a separate section. I first discuss the predictions based on information structure made by Katsos in 8.1, followed by the experiment itself in 8.2. Finally, in 8.3 I consider the relation to the main hypotheses of the current work based on the views of Van Kuppevelt and Van Rooij.

8. Katsos (2006) Experiment 4: Effects of information structure on SIs

8.1 Katsos' predictions based on information structure

Experiment 4 (Experiment 2 of Breheny et al. (2006)) is based on the claim of Katsos and colleagues that information structure properties of a sentence can be used to create either an SI-triggering or SI-blocking context. Their reasoning goes as follows: In English, sentences are typically constructed in such a way that the old information comes at the beginning, i.e. the topic of the sentence in terms of aboutness is usually sentence-initial (see also chapter 3 section 5.3). Therefore, an out-of-the-blue sentence is usually assumed to be 'about' whatever is in the earlier parts of the sentence. In Greek, which has flexible word order, this preference is even stronger. Therefore, (the Greek equivalent of) sentence (39) is more likely to be about the consultants than sentence (40), which is more likely to be about the director.

(39) Some of the consultants had a meeting with the director.

(40) The director had a meeting with some of the consultants.

So Katsos and colleagues argue that (39) is more likely to be related to issues like *What happened with the consultants?* or *Who did the consultants meet with?* than (40). They also claim that a context in which a contextual issue is raised about the set of consultants, is an SI-triggering environment for *some of the consultants*. They say:

> 'Where a question like *What happens with the Fs?* is accommodated into a context, that context would be an upper-bound context for the trigger 'Some of the Fs'. (Breheny et al. 2006, p. 446)

Therefore, even if there is no preceding context for (39) and (40), they predict that more SIs will be calculated for (39) than (40). Unfortunately, they do not provide off-line evidence to back up this claim.

8.2 Katsos' Experiment 4

As Katsos and colleagues realized the reading times of a critical segment in subject and object position are not directly comparable, they again looked at reading times of *the rest* in a following sentence. However, the fact that *the rest* is much closer to *some of the consultants* in the object condition than in the subject condition, introduces a bias. Therefore, Katsos introduced two baseline conditions which were similar to (39) and (40), but contained *only some* instead of *some*. This allowed him to compare the reading time difference on *the rest* between the *some* and *only some* conditions for both structural positions. Example items are given in (41)-(44):

(41) Subject *some*
Some of the consultants / had a meeting / with the director. / *The rest* / did not manage / to attend. /

(42) Subject *only some*
Only some of the consultants / had a meeting / with the director. / *The rest* / did not manage / to attend. /

(43) Object *some*
The director / had a meeting / with some of the consultants. / *The rest* / did not manage / to attend.

(44) Object *only some*
The director / had a meeting / with only some of the consultants. / *The rest* / did not manage / to attend.

Katsos and colleagues predicted that SIs will be calculated for *some of the consultants* in (41), but not in (43). Therefore, they predicted no reading time difference on *the rest* between (41) and (42). In both cases the facilitation should happen, in (41) due to the SI and in (42) due to the explicit *only some*. However, no facilitation of *the rest* is predicted for (43) as the SI is

not predicted to be calculated. Therefore, *the rest* should be read faster in (44) than in (43).

The results indeed show no difference in reading times of *the rest* in the subject conditions, but faster reading times of *the rest* in the object *only some* condition than in the object *some* condition. Katsos and colleagues take this to be evidence for the contextualist view. This is based on their idea that instead of an LB context, a scalar term in object position creates a *neutral* context, in which the defaultist view predicts the SI not be cancelled. As I pointed out before, I do not believe any context is truly neutral, as addressees will always make guesses about what came before in the context. Katsos and colleagues themselves also explicitly make this claim:

> 'However, as is well known, even single sentence utterances can create their own context through a variety of presupposition triggers and information-structure triggers.' (Breheny et al. 2006, p. 445)

Therefore, it is very strange that they consider a sentence like (40) as a neutral context, as it contains (at least) two presupposition triggers: the definite articles in *the director* and *the consultants*. These indicate these discourse referents were already present in the common ground when the sentence was uttered, and therefore the context can hardly be considered neutral.

Katsos and colleagues carried out an off-line control experiment to control for a possible bias of dispreferred continuation in the object *some* condition. In a continuation task, they found that there was no difference between the object *some* condition and the object *only some* condition in which NP (subject or object) was preferred as an antecedent for the subject of the next sentence, so this could not have been responsible for the reading time difference. In both conditions, around 50-60% of the times a continuation was chosen with a subject that referred back to as they put it: 'sets introduced by the final noun phrase' (Breheny et al. 2006, p. 451). Therefore, they concluded that the difference in reading times on *the rest* between these two conditions cannot be due to an effect of dispreferred continuation. It is however unfortunate that they only mentioned 'sets' and did not provide the information how many of these continuations were about the consultants that were present at the meeting and how many about the consultants that were not (the rest), as this would be an independent measure of how salient the complement set was, and therefore how likely it was that the SI was calculated. If as they claim no SIs were calculated in the object *some* condition, we would expect much fewer continuations about the complement set than in the object *only some* condition.

It would also have been interesting to conduct the same control experiment for the subject conditions, again looking at the difference between the set and the complement set continuations. This could have provided information whether the complement set is made as salient by an SI as by explicit *only*. Also, comparing the differences between the spontaneous continuations of all four conditions could provide support for the interpretation of the reading time data.

A weak point of Katsos' Experiment 4 is that two different structural positions are compared (albeit indirect by comparing the differences with *only some* in the same position). Therefore, it cannot be ruled out that the effect is due to the structural position itself, and not due to information structure. Perhaps for some unrelated reason *some of the* triggers more SIs in subject position than in object position. In the experiments presented the next chapter, I will therefore compare conditions in which the critical region is always in the same structural position, and information structure is manipulated explicitly by a question. However, first I will relate the results of Katsos' Experiment 4 to the hypothesis of focus-sensitivity of Van Kuppevelt and Van Rooij, as the data seem to go against this hypothesis.

8.3 Katsos topic-sensitivity versus Van Kuppevelt and Van Rooij's focus-sensitivity

Although Katsos and colleagues mention the theories of Van Kuppevelt and Van Rooij and refer to their ideas that contexts trigger contextual issues or questions, it seems their predictions are exact opposites of those of Van Kuppevelt and Van Rooij. Let us recapitulate Katsos' reasoning: he claims that because *some of the consultants* is topic in (39), it is likely to be old information in the context. Therefore, (39) is likely to be about the consultants, not the director, so (39) is likely to be an answer to (45) or (46):

(39) Some of the consultants had a meeting with the director.

(45) What happened with the consultants?

(46) Who did the consultants meet with?

Katsos claims a contextual question like (45) or (46) is an SI-triggering context for *some of the consultants*. So if (39) is given as an answer to (45) or (46), the SI will arise. Also, Katsos claims (40) is most likely an answer to (47) or (48).

(40) The director had a meeting with some of the consultants.

(47) What happened with the director?

(48) Who did the director meet with?

According to Katsos, (47) and (48) are not SI-triggering contexts for (40), as here *some of the consultants* is not old information. So in Katsos' reasoning, being old information in the context triggers SIs. This is the opposite of what Van Kuppevelt and Van Rooij claim, which is that SIs only arise if the scalar term is in the information focus, i.e. part of the *new* information. Relative to (47) and (48) the information focus of (40) is the VP or the direct object *some of the consultants*. Hence, Van Rooij and Van Kuppevelt predict the SI to arise there, contrary to Katsos' predictions. For (39) things are a bit more complicated.

Let us assume Katsos is right in claiming that (39) triggers a contextual question like (45) or (46), and (40) triggers (47) or (48). An important thing to notice is that in the contextual questions in (45) and (46) *the consultants* is used, while in the answer in (39) *some of the consultants* is used. So if indeed (39) triggers a question about the whole set of consultants (denoted by *the consultants*), *some of* is new information in the answer, not old information. However, it is not the part of the sentence that was questioned, as that is the VP (if (39) is an answer to (45)) or the direct object (if (39) is an answer to (46)). Relative to the contextual questions (45) and (46), the speaker changed something from the background or added something to the background of the question (from *the consultants* to *some of the consultants*). This indicates there is a *contrast* between the two. In that sense, relative to (45) and (46) *some of the consultants* in (39) is a *contrastive topic*.[5] So if Katsos and colleagues are right about the contextual questions that are triggered by the information structure of (39) and (40), they have

[5] The notion of *contrastive topic* is traditionally illustrated with the example in (i), due to Krifka (1991):

(i) Q: What did Bill's sisters do?
A: Bill's youngest sister kissed John.

In A's answer *Bill's youngest sister* is a contrastive topic. By picking a subset of the questioned topic, the answerer creates a contrast with Bill's other sisters that she could have picked as a topic. Similarly, if (39) is an answer to (45) as in (ii) below, the topic *some of the consultants* is contrasted with the other consultants.

(ii) (45) Q: What happened with the consultants?
(39) A: Some of the consultants had a meeting with the director.

been comparing the difference (in facilitation on *the rest*) between explicit *only some* and a scalar term in contrastive topic versus the difference between explicit *only some* and a scalar term in information focus.

If this is the case, there are two possible explanations for the difference found by Katsos, which are both compatible with Van Kuppevelt's and Van Rooij's predictions. First, it could be the case that contrastive topic triggers more SIs than information focus.[6] Especially as (39) and (40) are presented out of the blue, perhaps less SIs are calculated in information focus than when the sentence is embedded in a (explicit) context or preceded by an explicit question. Secondly, it could be the case that the SI is calculated in both cases (in subject position due to the contrastive topic and in object position due to information focus), but the complement set is made more salient in the subject case due to the contrast introduced by the contrastive topic. It is possible that a contrastive topic makes the complement set just as salient as with explicit *only some*, while the SI makes it less salient than with *only some*, which would explain why there was only a difference between *some* and *only some* in object position. Summing up, once we look more closely at the relation between the assumed contextual questions and the scalar sentences, we see that the data do not go against the focus-sensitivity of SIs as predicted by Van Kuppevelt and Van Rooij.

If we consider other questions that (39) and (40) could have been an answer to, ignoring Katsos and colleagues' claims about this, another possibility presents itself. If we consider only wh-questions, (39) could be an answer to the questions in (49) – (53):

(39) Some of the consultants had a meeting with the director.

(49) What did some of the consultants do?

(50) Who did some of the consultants have a meeting with?

(51) What did the consultants do?

(52) Who did the consultants have a meeting with?

(53) Who had a meeting with the director?

(51) and (52) are reformulations of resp. (46) and (45) of Katsos. What is crucial is that (49) and (50), where *some of the consultants* is subject, are

[6] For the relation between contrastive topic and SIs, see the work by Chungmin Lee, e.g. Lee (2006).

actually *infelicitous* in a context in which the set of consultants has been introduced, consider (54):

(54) context: On Monday, a team of consultants came to the office.
Q: # "What did some of them / some of the consultants do?"
Q: # "Who did some of them / some of the consultants have a meeting with?"

It is very strange to ask a question like this, where the subject is a non-specific set of consultants. Therefore, (49) and (50) are not possible contextual questions. We are left with (51) and (52), which as we saw above cause *some of the consultants* in (39) to be contrastive topic, and (53), relative to which *some of the consultants* is information focus. Crucially, there is no possible contextual question relative to which *some of the consultants* is (part of the) background, it is always new information. Now let us look at the possible contextual questions for (40), given in (55)-(58):

(40) The director had a meeting with some of the consultants.

(55) What did the director do?

(56) Who did the director have a meeting with?

(57) Who had a meeting with the consultants?

(58) Who had a meeting with some of the consultants?

(55) and (56) are reformulations of resp. (48) and (47) of Katsos, relative to which *some of the consultants* is information focus. Relative to (57) again something in the background is changed, so *some of the consultants* is contrastive topic.[7] Contrary to (49) and (50), it is felicitous to have *some of the consultants* as the direct object of the question in (58). Intuitively, (59) is much better than (54):

(59) context: On Monday, a team of consultants came to the office.
Q: "Who had a meeting with some of them / some of the consultants?"

[7] Depending on one's definition of topic, one might argue that *some of the consultants* is not topic here, but *the director* is. However, *some of the consultants* is contrastive with the questioned *the consultants* anyway.

Crucially, relative to (58), *some of the consultants* in (40) is background. So once we take more possible contextual questions into account, we see that *some of the consultants* in (39) is either information focus or contrastive topic, while in (40) *some of the consultants* is information focus, contrastive topic, or background. As Van Rooij & Van Kuppevelt predict SIs do not arise in the background, and the TVJT-experiments in chapter 3 showed SI-rates indeed go down in the background, it is not surprising that the condition on which the scalar term was possibly in the background, provided less facilitation of *the rest*. So also if we take a wider array of possible contextual questions into account, the results are in line with the predictions by Van Kuppevelt and Van Rooij.

The observation that questions with *some of the* in the subject like (49) and (50) are infelicitous, and therefore subject position seems to trigger more SIs than object position, extends to *or*. Sentence (60) below can only be an answer to (62), as (61) is infelicitous.[8]

(60) John or Bill talked to Mary.

(61) # Who did John or Bill talk to?

(62) Who talked to Mary?

As a result, *John or Bill* is always information focus, and the SI will arise. Contrastively, *John or Bill* in object position in (63) below is not always information focus, as (64) is fine (and so is (65)). This is why in the current work we are able to access the focus-sensitivity of the SI of *or* with sentences with *or* in object position.

(63) Mary talked to John or Bill.

(64) Who talked to John or Bill?

---

[8] Of course questions like (49) and (61) can be made felicitous, if we put *some of the consultants* or *John or Bill* in the common ground, e.g. (i) and (ii). However, it is unlikely that this will all be accommodated when e.g. (60) is interpreted out of the blue.

(i) A: You know about the consultants that were hired, right? I heard some of them had meetings with some exciting people.
B: Really? Who did some of the consultants meet with?

(ii) A: I heard John or Bill talked to someone famous.
B: Really? Who did John or Bill talk to?

(65) Who did Mary talk to?

This could be an explanation for the intuition that the SI of *or* in subject position seems to be much more robust than if *or* is in object position.

Notice that not every scalar term behaves this way. In chapter 7 we will test questions like (66), with the scalar term *most* in subject position:

(66) What did most students drink?

It seems the subject of a question can be non-specific to some extent, as in (66), but not too much, as in (49) and (61). I will not go into this issue any further, but leave it as a suggestion for further research.

I conclude that although the predictions of Experiment 4 of Katsos seem to be the exact opposite of the predictions following from the work of Van Kuppevelt and Van Rooij, the data can also be accounted for by their view. This holds if we take the contextual questions that Katsos assumed as well as if we consider a wider array of possible contextual questions that the target sentences were an answer to. I will now return to the current goal of testing the focus-sensitivity of SIs in an on-line experiment. In the next section I will discuss how the experiments of Katsos and colleagues serve as a starting point for the on-line experiments in the next chapter.

9. The experiments of Katsos (2006) as a starting point

The experiments of Katsos (2006) which I presented above, serve as a good starting point for the experiments that will be presented in the next chapter. If the slowdown they observed on the scalar region in UB contexts is indeed a reflex of SI-calculation, our hypothesis predicts we should be able to observe this slowdown if *or* is part of the focus compared to when it is not. In this section I will summarize some strong points and some weaknesses of the experiments of Katsos and colleagues, and explain some of the choices I made in the design of the reading time experiments presented in the next chapter.

Contrary to many earlier reading time experiments, Katsos and colleagues specifically considered effects of the wider context on scalar sentences. However, in all their experiments presented above, it was unclear which criteria were used to create the UB and LB contexts. It seems they were created based only on intuitions about UB-ness or LB-ness. Also, in different experiments different UB and LB contexts were used. In the experiments in the next chapter the contextual property of information focus

and the property of alternative relevance will be manipulated to create UB and LB contexts.

In Experiments 1, 3 and 5, the UB and LB contexts were not minimal pairs, as much material varied between conditions. Therefore, it cannot be ruled out that the differences that were observed between the conditions were not due to the inference under investigation, but were caused by other factors. In order to avoid this in my experiments, contexts will be used that only differ on the property under investigation. In Experiments 1 and 3 of Katsos and colleagues the target sentence that the critical region was part of varied between conditions, making a comparison between conditions harder. In the experiments that I will present in chapter 6, I will use identical target sentences across conditions. Differences between the conditions will then necessarily reflect an effect of the wider context.

In all of Katsos' experiments in which there was a direct comparison between UB and LB contexts (so all experiments except Experiment 4), the contextual manipulations were problematic: In Experiment 1, the exclusivity was already given in the UB contexts, in Experiment 5 the UB contexts contained the stronger scalar term, and in Experiment 1, 2 and 3 the target sentences in the LB contexts contained a modal verb. In my experiments I will vary only the contextual properties under investigation, thereby avoiding these biasing factors.

No off-line data were reported for the items in Katsos' Experiments 1, 4 and 5 to check whether the LB contexts were truly SI-blocking, and the UB contexts were truly SI-triggering. For instance, the LB contexts of Experiment 5 did not strike me as truly SI-blocking, and the assumptions about how information structure affects SIs in Experiment 5 also calls for off-line support. I will use the same or very similar items in the reading time experiments as I used in the off-line experiments presented in the previous chapters.

In Experiment 3 and 5 of Katsos and colleagues found a smart way to assess the interpretation of the sentence and measure reading times on the scalar region itself at the same time. They achieved this by measuring reading times on a region in the next sentence which was facilitated by the SI. As this region was read faster as the scalar region was read slower, this supported the idea that the slower reading times on the scalar region were due to SI-calculation. I will adopt this methodology in the second experiment presented in the next chapter, but instead of using following material that is compatible with the SI, I will use material that is incompatible with it. So instead of facilitation I will be considering a possible penalty caused by an interpretation with an SI. The advantage of this approach is that it allows for a comparison with a baseline condition with the top-of-the-scale item *and*, which should not trigger this penalty.

Unfortunately, Katsos and colleagues did not include such control conditions in which the scalar term was replaced by a non-scalar or a top-of-the-scale scalar item, which would provide further evidence for the delay being due to the SI. Finally, I will try to avoid some of the methodological problems of the experiments of Katsos, by avoiding sentence-final position of the critical region, and stricter constraints on the length (in syllables) of the critical region and spillover regions. Now that we know what to look for and what to avoid, it is time to present the experiments.

CHAPTER 6: FOCUS-SENSITIVITY OF SCALAR IMPLICATURES IN REAL-TIME PROCESSING

1. Introduction

The TVJT experiments in chapter 3 showed that more SIs are calculated if the scalar term *or* is part of the information focus of the sentence than if it is part of the background. However, with the TVJT paradigm the participant already knows what the actual situation is when she interprets the sentence containing the scalar term, which makes it unlikely that she will calculate an SI. In chapter 4 I avoided this problem by using the PWJT paradigm, in which the situation was no longer given in the story. Participants were asked to judge whether the situation which could have been described by a sentence with the stronger term *and* was possible. However, this experimental question possibly overruled the critical manipulations of focus and alternative relevance, as the question itself might have become the issue relative to which the target sentence was interpreted. Therefore, I concluded we need a paradigm in which the actual situation can be left out and we do not need an experimental question to assess the SI. In the previous chapter I considered a number of processing studies from the literature which suggest measuring reading times might be a suitable method to test our hypothesis. The results of the experiments by Katsos (2006) suggest that SI-calculation brings along a processing cost that is reflected by longer reading times on the scalar region. Therefore, the critical prediction of the experiments in this chapter is that if a scalar term is in the information focus of a sentence it will be read slower than if it is part of the background.

In this chapter I present two experiments that test this prediction, Experiments 8 and 9. Both experiments were conducted with the self-paced reading (SPR) paradigm. In this paradigm, participants read a text bit by bit from a computer screen and progress through the text by pressing a button. Reading times are measured as the time between two button presses. This is a widely adopted paradigm in the language processing literature, and the experiments of Katsos (2006) suggest it can be used to reveal the processing cost of SI-calculation. In both experiments presented in this chapter, information focus was manipulated in the same way as in previous experiments (except Experiment 3), by varying the explicit question which the target sentence was an answer to.

In Experiments 8 and 9 I compared reading times of constituents that were part of the focus to those of constituents that were part of the background. However, some experimental studies in the literature have produced results that suggest focus in general might affect reading times,

irrespective of SIs. Therefore, we need to take these results into account in the current experiments. In section 2 I briefly address these results and I argue that a possible independent effect of focus on reading times can be controlled for by including suitable control conditions.

I present Experiment 8 in section 3. In this experiment I used stories similar to those of (TVJT) Experiment 2, but with the actual situation removed from the story. Similar to Experiment 2, only the explicit question which determined the focus structure of the target sentence varied between conditions. Two control conditions with *and* were included to test whether an effect was actually due to SI-calculation. These control conditions were also needed to control for the repetition of the critical region (*A or B*) from the question in the non-focus condition.

Experiment 8 did not reveal a slowdown in reading on the scalar region in the focus condition. I will argue this is due to the very minimal nature of the contexts used in the items, as a result of which there was no incentive for the participants to calculate the SI. To correct this I used the richer contexts of the PWJT-experiments in Experiment 9, presented in section 4. These contexts contained a conditional in which the alternative of the stronger scalar alternative was manipulated ('alternative relevance', see section 2 of chapter 4). This made the SI relevant for the outcome of the story and hence participants had an incentive to calculate it. In Experiment 9 I also added an additional reading time measure of SIs. I included a continuation sentence in which a plural pronoun (*them*) was used to refer back to the *A or B* constituent. As a plural pronoun is incompatible with an interpretation *A or B* with SI (*A or B but not both*), but not with a reading without SI (*A or B or both*), a delay in reading suggests the SI was calculated.

The main goal of these experiments is to find further evidence for the effect of information focus on SIs, controlling for the shortcomings of the TVJT and the PWJT. I will discuss the implications of the results for the hypothesis of focus-sensitivity of SIs in section 5.1. However, the experiments also address the question whether SIs are actually calculated on-line during reading and whether this leads to an observable delay, as was claimed by Katsos (2006) i.a. Especially Experiment 9, in which an independent measure of SI is included (the delay on a region in the next sentence) can provide support for the view that SI-calculation is costly, if this delay co-occurs with a delay on the scalar region. In section 5.2 I will consider the implications of the current results for this question and the implications for the defaultist/contextualist debate.

2. Previous studies on the effects of focus on reading times

A number of studies in the psycholinguistic literature indicate that focus affects reading times. As I pointed out in chapter 3, the notion of *focus* in these papers is a bit different from the way I am using it in this work. To avoid confusion, I will again use the term *emfocus* for the psycholinguistics notion of focus, defined by Birch & Rayner (1997) as: 'the information that is newly asserted in a discourse, sometimes contrastive, and is most prominent or emphasized within the sentence' (p. 653). See chapter 3 section 7.1 for discussion.

Birch & Rayner (1997) found that emfocus affects reading times. Using eye-tracking, they compared reading times on regions of a sentence that were either questioned or not. For instance, they compared reading times of *cards* in (1) and (2) (my italics):

(1) emfocus condition
Q: What were the soldiers playing?
A: The soldiers in the underground bunker were playing *cards* to relieve their boredom.

(2) non-emfocus condition
Q: Where were the soldiers?
A: The soldiers in the underground bunker were playing *cards* to relieve their boredom.

Cards is questioned in (1) and therefore has emfocus, but not in (2). Birch & Rayner found no difference in first-pass reading times, but did find significantly longer second-pass reading times on the critical region in the emfocus condition than in the non-emfocus condition.[1] When they used a longer critical region, as in (3) and (4), first-pass reading times were significantly longer on the emfocused constituent, but second-pass reading times were not.

(3) emfocus condition long region
Q: Where were the soldiers?
A: The soldiers *in the underground bunker* were playing cards to relieve their boredom.

[1] First-pass reading time is the time spent on the region before the participant either moves on or looks back. Second-pass reading time is the sum of refixations on the region after it has been left for the first time (see Koornneef 2008).

(4) non-emfocus condition long region
Q: What were the soldiers playing?
A: The soldiers *in the underground bunker* were playing cards to relieve their boredom.

I think none of the target sentences above are felicitous answers to the questions. Especially in (2) and (3), it seems highly marked to embed the answer to the question in a relative clause. Also, the fact that a lot of material is added to or changed from the background in all conditions, makes these results difficult to interpret. For instance, in all conditions *The soldiers in the underground bunker* could be considered a contrastive topic, instead of information focus or background (see 8.3 of chapter 5). If we ignore these problems, the results indicate that questioned material is read slower in an answer than non-questioned material, which is important for the current enterprise.

However, another study shows the opposite pattern. Morris & Folk (1998) manipulated emfocus with *it*-clefts, and found that an emfocused constituent was read faster than a non-emfocused constituent. For instance, they found that *accountant* was read faster in (5) than in (6):

(5) emfocus condition
While the waiter watched, it was the *accountant* who balanced the ledger a second time.

(6) non-emfocus condition
It was the waiter who watched while the *accountant* balanced the ledger a second time.

Where the study by Birch & Rayner found slower reading times on an emfocused constituent and Morris & Folk found faster reading times on an emfocused constituent, Ward & Sturt (2007) did not find a difference in reading time at all. They observed no difference in reading time of *exit* in (7) and (8), where emfocus was manipulated with an embedded question.[2]

(7) emfocus condition
I couldn't decide which seat to take at the theatre.
I hoped the seat by the *exit* would give me a good view.

[2] Ward & Sturt conducted a change detection task. I report the data of the first display of the target sentence here.

(8) non-emfocus condition
I couldn't decide whether I liked the new theatre layout.
I hoped the seat by the *exit* would give me a good view.

It is hard to interpret the implications of these data for the more restricted notion of information focus investigated here. The fact that the evidence is not converging also makes it hard to predict how an effect would interact with a delay due to SI-computation. However, what the results of these experiments show is that focus might in some way affect reading times. Therefore, I included control conditions in Experiments 8 and 9 that control for effects of focus that are unrelated to the inference under investigation.

3. Experiment 8: Testing the focus-sensitivity of SIs using reading times

3.1 Setup and items

One of the problems of the experiments of Katsos (2006) is that the UB (SI-triggering) contexts and LB (SI-blocking) contexts were not minimal pairs, but varied on several points from each other, which could have affected the reading times of the critical region. In order to keep the conditions as minimally different from each other as possible, I used the items of (TVJT) Experiment 2 (presented in chapter 3) as a starting point. In that experiment the only thing that varied between conditions was the explicit question, which was the critical manipulation of information focus. I repeat the example items of Experiment 2 here in (9) and (10), with the differences between them given in boldface:[3]

(9) Experiment 2 focus condition
Julie and Karin were searching for marine animals on the beach. After some searching Julie found a crab. Not much later she also found a starfish. Unfortunately, Karin didn't find anything.
When **Karin** returned, her mother asked **what kind of marine animals Julie had found**.
Karin answered that Julie had found a crab or a starfish.

[3] Another difference between the two conditions was *When Karin returned / When they returned*. See fn. 6 of chapter 3 for discussion.

(10) Experiment 2 non-focus condition
Julie and Karin were searching for marine animals on the beach. After some searching Julie found a crab. Not much later she also found a starfish. Unfortunately, Karin didn't find anything.
When **they** returned, their mother asked **who had found a crab or a starfish**.
Karin answered that Julie had found a crab or a starfish.

In (9) the information focus of the target sentence is *a crab or a starfish*. The scalar term *or* is in this constituent, so the SI is predicted to arise. In (10), *Julie* is the information focus of the sentence, so the scalar term is part of the background and the SI is not predicted to arise. Indeed, more SIs were observed for items like (9) than items like (10) in Experiment 2, reflected by more 'false' judgments for the answer in (9) than in (10).

I adapted these items by taking out the information about what actually happened (who found what), to avoid the discrepancy between story and target sentence which was problematic in the TVJT experiments. Example items of Experiment 8 are (11) and (12), with differences marked in boldface.

(11) Condition 1: focus *or*
Hielke and Sietse were out fishing on their boat.
When **Hielke** returned, their father asked **what kind of fish Sietse had caught**.
Hielke answered that Sietse had caught a pike or a carp.
Father said that Sietse was a lucky devil.

(12) Condition 2: non-focus *or*
Hielke and Sietse were out fishing on their boat.
When **they** returned, their father asked **who of them had caught a pike or a carp**.
Hielke answered that Sietse had caught a pike or a carp.
Father said that Sietse was a lucky devil.

A number of items of Experiments 1 and 2 was adapted to the new format, and a number of new items were added.[4] [5] The items also featured a

[4] Although Experiment 8 was based on the items of Experiment 2, the stories of Experiment 1 also provided a good baseline for items of Experiment 8 after the extra material (e.g. the conditionals) was taken out.

continuation sentence (the last sentence in (11) and (12)) in order to be able to measure reading times on the final region of the target sentence. 20 different verbs were used, 3 of which occurred more than once: *bought* (3x), *eaten* (2x) and *made* (2x). Only verbs and objects were used for which it was possible and plausible that the person 'verbed' both, avoiding that the exclusivity of *or* already followed from the story, as in Katsos (2006) Experiment 1 which contained VPs like *was born in A or B*. Based on the results of Katsos (2006), and the off-line results of the TVJT-experiments, we expect the region containing *or* (*a pike or a carp*) to be read slower in (11) than in (12).

Another point I raised about the experiments of Katsos (2006) was that there was no check whether the delay on the scalar term was due to an SI. In Experiment 8, I included two control conditions, which were identical to the two test conditions but in which *or* was replaced by *and*, see (13) and (14), where for easy reference *and* is marked in boldface (not in the actual items).

(13) Condition 3: focus *and*
Hielke and Sietse were out fishing on their boat.
When Hielke returned, their father asked what kind of fish Sietse had caught.
Hielke answered that Sietse had caught a pike **and** a carp.
Father said that Sietse was a lucky devil.

(14) Condition 4: non-focus *and*
Hielke and Sietse were out fishing on their boat.
When they returned, their father asked who of them had caught a pike **and** a carp.
Hielke answered that Sietse had caught a pike **and** a carp.
Father said that Sietse was a lucky devil.

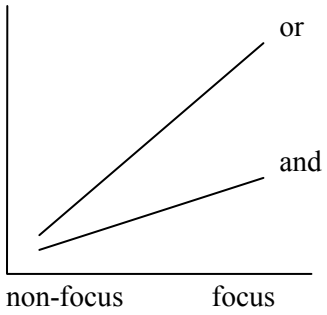
[5] Not all of the items of Experiment 1 and 2 were suitable for a reading experiment, as in order for the question in the non-focus condition to be somewhat natural, the objects in *A or B* could not be too random or specific. Compare for instance (12) to (10): it is natural to ask who caught a pike or a carp, as both are kinds of fish which occur frequently in the 'fishing script'. However, it is less straightforward why anyone would ask specifically about a crab or a starfish, as in (12). This kind of unnaturalness is not problematic in an off-line task, in which there is no time pressure and the participant can reread the story as often as she wants, but it might introduce more experimental noise in an on-line task in which reading times are measured and participants have to memorize parts of the story. Therefore, I tried to use somewhat natural pairs of objects in Experiment 8.

As *and* is the top item of the <or, and> scale, it will not trigger SIs in UE episodic sentences as the target sentences in (13) and (14). Therefore, any difference between the *or*-conditions that is due to SI-calculation should not be observed in these control conditions.

There are more reasons why these two control conditions were necessary. As a result of the focus manipulation, the critical region *A or B* was repeated from the question in the non-focus condition, while it was new in the focus condition. This was unavoidable, as in order for *A or B* to be old information in the non-focus condition, it had to be present in the question (and absent in the focus condition to be new information). However, this repetition could lead to faster reading times on the critical region in the non-focus condition. Therefore, we have to compare the reading times of the experimental conditions to the reading times of the *and*-conditions, which also contain the difference in repetition. The control conditions with *and* also control for a possible effect of focus on reading times in general, discussed in section 2 above.

So instead of considering an absolute difference on reading times of *A or B* in the focus and the non-focus conditions, we have to consider these reading times relative to their baseline conditions with *and*. Crucially, we predict an interaction between the factors focus (focus vs. non-focus) and connective type (*and* vs. *or*): Due to the repetition of the critical region from the question in the non-focus conditions, we expect both focus conditions (*or* and *and*, C1 and C3) to be read slower than their non-focus counterparts (resp. C2 and C4). However, the difference between the two *or*-conditions is predicted to be bigger than the difference between the two *and*-conditions, due to the processing cost of the SI in the focus *or*-condition. (Equivalently, the difference between the two focus conditions is predicted to be bigger than the difference between the two non-focus conditions (if any)). The predicted pattern looks like Figure 1:

Figure 1: predicted reading time pattern Experiment 8



3.2 Critical region and spillover regions

Another point I raised about the experiments of Katsos (2006) in the previous chapter (and which is a general principle in SPR-studies) is that it should be avoided that the critical region is in sentence-final position. This is problematic as the reading time for the critical region will then be contaminated with sentence wrap-up effects. Another reason for having material follow the critical region in the sentence is that we can use that material as spillover regions, in case the effect is delayed. Fortunately, Dutch has SOV verb order in a subordinate clause, so the auxiliary and the verb follow the direct object in our target sentences and they can serve as spillover regions. Example item (15) is translated literally from Dutch. In this example slashes indicate region breaks and double slashes indicate line breaks.

(15) example segmentation Experiment 8
Hielke and Sietse/were out/fishing/on their boat./When they/returned,//
their father asked/what kind of fish/Sietse/had caught.//
Hielke answered/that Sietse/a pike or a carp/had/caught./Father/said //
that Sietse/was a lucky devil.//

The critical region (CR) (*a pike or a carp* in (15)) was always the third region of the third line, the 13th region in total. It mostly contained count nouns with the indefinite determiner 'een' (*a*), although a few items contained names or mass nouns (without determiners). With one exception, the critical region consisted of 3 to 5 words.[6] The first spillover region (SO1) was the auxiliary, which was identical in all items (*had*). The second spillover region (SO2) was the (past participle) verb, which was always identical to the verb form used in the question. As reading times on the last region of a line are unreliable due to (planning of) eye-movements to the next line, I made sure the SO2 region was not the last region of the line, by including two regions of the final sentence on the same line (*Father / said* in (15)).

3.3 Design

24 story quadruples like (11)-(14) were created and divided over 4 lists in a Latin-square design. The items are given in Appendix 6. Every list contained

[6] The one exception was the critical region *a bag of (potato) chips or a bag of candy*, which in Dutch contains 7 words ('een zak chips of een zak snoep').

6 items per condition. The lists also contained 30 fillers (the same in all lists) which were similar to the test items, but did not contain *A or B*.[7] The fillers did feature other complex NPs like *a purple dress with dots*. Otherwise, the fillers had the same structure as the test items: an introduction followed by a question-answer dialogue in indirect speech and a final remark by one of the speakers. Typical *who*-questions (*Who started the fight?*) as well as typical *what*-questions (*What did he get for his birthday?*) appeared in the fillers.

To ensure the participants read the items for comprehension one third of the items (both test items and fillers) were followed by a verification statement about different parts of the stories, which had to be judged true or false by the participants. Half of these were true and half were false. The 4 lists were distributed evenly over participants, and the items were presented in a newly randomized order for every participant, through a randomization procedure in the experimental software.

3.4 Participants and procedure

46 native speakers of Dutch participated in the experiment. They were all students of Utrecht University. None of them had any prior knowledge of the topic and none of them had reading problems. All participants were paid for their participation. The experiment was run on computers in sound proof booths in the lab, using E-Prime software (Psychology Software Tools). The *non-cumulative moving-window* self-paced reading paradigm was used. In this paradigm, all the words of the story are replaced by lines (strings of underscores) representing their lengths and positions in the story, and one segment (consisting of a word or a couple of words) at a time is shown at the participant's button press. With the next button press the segment is replaced by lines again (hence *non-cumulative*), and the next segment is shown. This way participants read through a text one segment at a time, but the reading process is more natural than when words are presented in the center of the screen (which is called the *fixed*-window paradigm). The periods indicating the end of the sentences in the stories were also visible throughout, so participants could anticipate the end of a sentence. Participants progressed through the text by pressing the space bar of the keyboard. The experimental software registered the elapsed time between two presses of the space bar, which represented the reading times of the regions.

[7] Actually, two fillers did contain *A or B*, but not as a direct object in an episodic sentence: *he wanted to paint the room purple or orange* and *she wanted to adopt a baby from Senegal or Ivory Coast*.

As the participant pressed the space bar after reading the last region of an item, the story disappeared and either a verification statement about the story appeared on the screen, or an instruction to press the TAB key to continue to the next item. The verification statements had to be judged true or false by pressing the '1' (true) or the '2' (false) key on the keyboard. After the participant pressed the '1' or '2' key or the TAB key, the next item appeared on the screen. Participants were instructed before the experiment and were presented with 3 practice trials, 2 of which were followed by a verification statement. They were told to read at a normal pace but to make sure that they read all the material in a segment before moving on. Participants were told to press the space bar with their dominant hand, and not switch hands during the experiment. Although reading speed varied considerably between participants, most participants completed the experiment in around 20 minutes.

3.5 Results

Two participants were excluded from the analysis because they answered less than 75% of the verification statements correctly (they scored resp. 72% and 56% correct). The reading times on the segments of the target sentence of the remaining 44 participants were analyzed. The effect of outliers was controlled as follows: reading times under 200 ms. were removed from the data (1% of the data), and reading times more than 2.5 SD higher than the grand mean for that region were smoothed to that value (2.8% of the total data set). Smoothing was chosen instead of exclusion as these long reading times might include SI-computation, but we do not want the effect of extremely long reading times on the analysis to be too big. I will discuss the results for the critical region and the rest of the sentence separately.

3.5.1 Results *A or B/A and B* region

The smoothing point for the critical region was 2560 ms, 2.9% of the data of the critical region was smoothed to this value. The average reading times (calculated over all participants and test items) are given in Table 1:

Table 1: average reading times *A or B/A and B* region Experiment 8

	reading time (ms)
C1 (focus or)	981
C2 (non-foc or)	899
C3 (focus and)	978
C4 (non-foc and)	857

2x2 ANOVAs over subjects (F1) and items (F2) were computed with factors focus and connective type as within-subjects and within-items factors. The analysis showed a significant main effect of focus: $F1(1,43) = 12.49$, $p = 0.001$, $F2 (1,23) = 11.23$, $p = 0.003$, which was expected due to the repetition difference. No main effect of connective was observed (F1 and F2 both < 1). Crucially, there was no significant interaction of focus and connective type: F1 and F2 both < 1.

Paired t-tests over subjects (t1) and items (t2) were calculated between pairs of conditions that were minimally different (varied only on focus or only on connective), giving four comparisons. I report two p-values for every comparison, one uncorrected and one corrected for multiple (4) comparisons (Bonferroni correction). All p-values are two-tailed, unless stated otherwise. The t-tests revealed that for both *or* and *and*, the focus condition was read slower than the non-focus condition, although the difference for *or* was only significant without correction and marginally significant in the items analysis: *Or*-conditions: C1-C2: $t1(43) = -3.377$, $p = 0.044$ (no correction) / 0.176 (Bonf. corrected for 4 comp.), $t2(23) = -1.927$, $p = 0.066/0.264$. *And*-conditions: C3-C4: $t1(43) = -3.377$, $p = 0.002/0.008$, $t2(23) = -3.423$, $p = 0.002/0.008$. Comparison of *or* and *and* did not return a significant difference in the focus conditions (C1-C3: both t's < 1) nor in the non-focus conditions (C2-C4: $t1(43) = 1.079$, $p = 0.287/1.000$ and $t2(23) = 1.419$, $p = 0.169/0.676$). We can conclude the predicted effect of SI-computation was not observed.

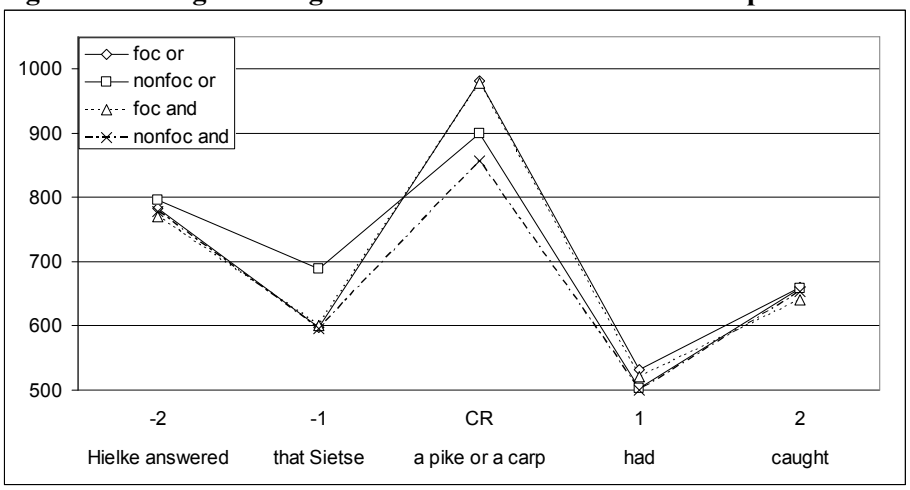
3.5.2 Reading times whole *A or B / A and B* sentence

The average reading times (calculated over all participants and all test items) of the segments in the target sentence are given in Table 2 and Figure 2:

Table 2: average reading times *A or B/A and B* sentence Experiment 8

	-2	-1	CR	SO1	SO2
	Hielke answered	that Sietse	a pike and/or a carp	had	caught
C1 (foc or)	783	597	981	532	660
C2 (nonfoc or)	796	688	899	503	658
C3 (foc and)	770	601	978	522	641
C4 (nonfoc and)	777	596	857	500	654

Figure 2: average reading times *A or B/A and B* sentence Experiment 8



The main effect of focus spilled over to SO1: F1(1,43) = 6.09, p = 0.018, F2(1,23) = 7.46, p = 0.012. Paired t-tests revealed that the difference between C1 and C2 in this region was significant in the subjects analysis without correction: t1(43) = -2.025, p = 0.049/0.196, t2(23) = -1.625, p = 0.112/0.448, but the difference between C3 and C4 was not: t1(43) = -1.690, p = 0.098/0.392, t2(23) = -1.577, p = 0.128/0.512. Other than that, no significant differences were observed in the spillover regions.

The regions preceding the critical region showed an unexpected result. No significant differences were observed in the −2 region, but in the −1 region the non-focus *or*-condition (C2) was read almost 100 ms slower than the other conditions. As C1 and C3 are identical at that point (and their reading times did not differ significantly), I compared C2 to the pooled data of C1+3 and to C4. These comparisons showed the differences were highly significant: C2 vs. C1+3: t1(43) = -3.253, p = 0.002/0.008, t2(23) = -3.027, p = 0.006/0.024, C2-C4: t1(43) = 3.577, p = 0.001/0.004, t2(23) = 3.632, p = 0.001/0.004. It is not surprising that the non-focus *or*-condition (C2) was

read slower than the focus *or*-condition (C1) and the focus *and*-condition (C3) (which were at that point identical). In the non-focus *or* condition, the segment containing the subject was actually the answer (so the information focus). In line with the results by Birch & Rayner (1997) discussed in section 2, constituents which are the information focus might be read slower than when they are background. Also, in the focus condition the name was repeated from the question, which might have sped up the reading of it in the answer. The surprising finding is that the non-focus *or*-condition (C2) was also read slower than the non-focus *and*-condition (C4). I will briefly consider a possible explanation here, but I will argue this observation in the −1 region cannot explain the absence of the predicted effect in the critical region.

At the -1 region the only difference between C2 and C4 is the use of *or* versus *and* in the question, so the difference has to be caused by the use of *or* and *and* in the question. A possible explanation for it is that the questions might trigger different presuppositions. Perhaps *Who verbed A and B?* more readily triggers the presupposition that someone verbed A and B than *Who verbed A or B?* triggers the presupposition that someone verbed A or B, due to the less specific nature of the latter question. This difference in presuppositions could explain the reading time difference. In the non-focus *and*-condition, the presupposition that someone verbed A and B can already be accommodated at the question, while in the non-focus *or*-condition, the presupposition that someone verbed A or B has to be accommodated at the answer, possibly causing a delay. In other words, during the interpretation of the answer in the non-focus *and*-condition participants only have to fill in the value of the variable in the presupposition, while in the non-focus *or*-condition participants have to accommodate the presupposition that indeed someone verbed A or B, and fill in the value of the variable.

This difference in presuppositions is just a hunch about how the unexpected data point might have come about, and I will leave it as a suggestion for further research. However, what is important for our current purposes is that this unexpected observation in the −1 region cannot account for the fact that the predicted effect in the critical region was not observed. Irrespective of whether there was an unexpected effect in the −1 region in either of the non-focus conditions that might have spilled over to the critical region, the focus conditions did not show the difference (between *and* and *or*) in the critical regions which was expected based on SI-calculation. If SIs were calculated in the focus *or* condition, this should have led to a delay relative to the focus *and* condition, which was not observed. Therefore, the explanation of the effect is not relevant for the implications of this experiment for the hypothesis of focus-sensitivity of SIs.

3.6 Discussion

The results of the critical region did not show the predicted interaction of focus and connective type. The question arises whether the experiment was sensitive enough to measure the effect. However, we did observe the effect of repetition on the critical region, leading to faster reading times in the non-focus conditions. This indicates the experiment was sensitive enough to show an effect of the focus manipulation, and that participants did actually read the part that contained the critical manipulation (the question). The fact that the difference between the two *or*-conditions was not bigger than between the two control conditions with *and*, could mean (at least) two things: First, it could mean that SI-computation does not lead to an observable delay in reading times. This goes against the findings of Katsos and colleagues, discussed in the previous chapter. Second, it could mean that participants did not calculate more SIs in the focus *or*-condition than the other conditions. This goes against the results of the TVJT experiments in chapter 3. As the items in this experiment were derived from those of (TVJT) Experiment 2, this seems implausible. However, let us look at the differences between the two experiments in more detail.

The first difference is that the items of the TVJT contained the actual situation, which could have also been described by *A and B*. I pointed out why this is problematic in chapter 3, so in this experiment this situation was taken out. The second difference is that in the TVJT, participants had to judge the target sentence relative to the *A and B* situation, while in this experiment participants merely had to read the sentence and memorize it. The result of these changes was that in the current experiment, there was no reason for the participants to calculate the SI anymore. In the TVJT experiments, at the point where participants read the target sentence they had already read the *A and B* situation, and they knew they had to judge the sentence relative to that situation. For this task the SI was crucial, as the truth-value of the sentence depended on it. Hence, there was a reason to calculate the SI. However, in the reading experiment the *A and B* situation was taken out and participants did not have to judge the target sentence relative to it. Participants only had to read the sentence and memorize the information in order to be able to answer the verification statements. The verification statement was never about *A or B* versus *A and B*, as I wanted to avoid reference to the critical inference. Therefore, nothing hinged on the SI anymore and there was no reason for the participants to calculate it. Hence, in Experiment 8 there was a crucial difference between the hearer in the story and the participant in the experiment. Perhaps if the participant would be in a dialogue as in the focus condition of Experiment 8 in real life, she would calculate the SI, as her goal would be to find out what the world is

like. However, as the goal of Experiment 8 is merely reading and memorizing a story, there is no reason to calculate the SI. Remember that the contextualists claim SIs are costly inferences, which will only be drawn if they are worth the effort. On this view, it would not be surprising if participants did not calculate SIs in this experiment. This could explain the absence of an effect of SI-calculation in the focus *or*-condition.

The methodological problem we are facing is the following: there should be an incentive for the participant to calculate the SI. However, we want to avoid explicitly providing the actual situation as in the TVJT experiments, as this is not in line with how SIs arise in actual conversation and it introduces the infelicity of using a weaker term. We also do not want to explicitly ask about the *A and B* situation as in the PWJT experiments, as that will overrule the contextual manipulations. However, remember that in the PWJT experiments we also manipulated alternative relevance. We can use this manipulation to increase the relevance of the SI for the participants. In the PWJT experiments, we used items like (16):

(16) Experiment 5 [+alternative relevance, +focus] condition
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found at least two marine animals on the beach, that person got to stay up late that night.
The friend said: "Oh, and what had Laura found on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

The requirement of finding at least two animals makes the alternative with *and* relevant in the story. In (16), the friend asked the question to find out whether the conditional was satisfied, and it was important for the conditional whether Laura found both or only one of the two animals. This contextual manipulation also provided a reason for participants to calculate the SI, as it has consequences for the outcome of the story. This is contrary to the items of Experiment 8, where there were no consequences that depended on the SI. Even when the task of participants is merely to read a story like (16) and memorize the information, there is an incentive to calculate the SI as it is relevant in the story. Therefore, I decided to do a follow-up reading experiment with the stories of the PWJT experiments in which conditionals like this were used. However, in chapter 4 I already indicated that it is impossible to create coherent contexts in which a [+ alternative relevance] conditional like in (16) is continued by a [-focus] type question (like *Who found a crab or a mussel?*). Therefore, I contrasted the

[+alternative relevance, +focus] items with the [-alternative relevance, -focus] items of Experiment 5. This means the follow-up experiment tested the cumulative effect of focus and alternative relevance.

In the follow-up experiment, I wanted to include another measure of whether more SIs were calculated in the focus condition than in the non-focus condition, as Experiment 8 raised some doubt about whether this can be observed in reading times on the *A or B* region. In chapter 5 we saw that Katsos (2006) found an elegant way of doing this. He measured reading times on a subsequent region which was facilitated by the SI. I decided to use a following region that is incompatible with the SI: a plural pronoun referring back to *A or B*. If the SI is calculated for *A or B*, this continuation will be unexpected and should lead to a delay in reading. Crucially, this continuation is expected for *A and B*, so again control conditions with *A and B* will provide a good baseline. In this setup, we have two indicators of SIs: longer reading times on the scalar region and longer reading times on the following incompatible region. The reading times on the following region can serve as a check for the effect of SIs on reading times on the *A or B* region.

4. Experiment 9: Testing the effect of focus and alternative relevance on SIs using reading times

4.1 Setup and items

I used the items of (PWJT) Experiment 5 in Experiment 9. Examples of the [+alternative relevance, +focus] condition and the [-alternative relevance, -focus] condition of Experiment 5 are repeated in (17) and (18), with the differences given in boldface. For easy reference, I will call the [+alternative relevance, +focus] condition the *FocRel* condition and the [–alternative relevance, -focus] condition the *non-FocRel* condition. Differences between the conditions are again given in boldface (not in the actual items).

(17) Experiment 5 FocRel condition
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **at least two marine animals** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **what had Laura found** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(18) Experiment 5 non-FocRel condition
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **a crab or a mussel** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **who of Laura and Barbara had found a crab or a mussel** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

In (17) the conditional *if one of them had found at least two marine animals* triggers alternative relevance, and the question *what had Laura found?* causes a *crab or a mussel* to be the information focus of the answer. Therefore, we expect the SI to be calculated in this condition. In (18) it does not matter for the reward whether Laura found both a crab and a starfish or only one of the two, so for this story the stronger alternative was not relevant. The *who*-question causes *Laura* to be the information focus of the target sentence, and *a crab or a mussel* to be the background. Thus, this condition has neither of the contextual properties considered crucial for SIs and no effect of SI-calculation is expected.

16 items were taken from Experiment 5, with minor changes to satisfy number of syllables requirements, and 20 new items were created. The target sentences contained 23 different verbs, 6 of which were used more than once: *bought* (5x), *made* (3x), *taken* (2x), *found* (2x), *brought* (4x). Again, I made sure it was possible and plausible that both objects were 'verbed' in all items. A continuation sentence was added to the stories in which either a singular or a plural pronoun was used to refer back to *A or B*. For (17) and (18), the continuation sentences were (19) and (20), with the pronouns marked in boldface (not in the actual items):

(19) plural: Laura said later that she had found **them** in the surf near the lighthouse.

(20) singular: Laura said later that she had found **it** in the surf near the lighthouse.

If participants calculated the SI *Laura did not find both a crab and a mussel* in (17) or (18), the plural pronoun in (19) is unexpected. In that case it will be harder to find an antecedent for the pronoun (the SI has to be cancelled) than if the SI is not calculated (*Laura found a crab or a mussel and possibly both*). Therefore, the FocRel condition should lead to longer reading times on the plural pronoun than the non-FocRel condition. The singular pronoun is a possible continuation of the sentence irrespective of the SI, so no difference is expected there.

The speaker switch between the *A or B* sentence and the continuation is necessary as *A or B* potentially triggers an Ignorance Implicature that the speaker does not know exactly what Laura found (see chapter 8). It would then be strange for the speaker to refer back to *A or B* with a pronoun in the next sentence, as this would imply she does know exactly what Laura found. However, I avoided this problem by letting the person who did the searching, and therefore knows what she found, use the pronoun.

The Dutch pronoun in the plural continuation (the *them*-conditions) was 'ze', which is ambiguous between third person plural nom./acc. (*they/them*) and third person singular feminine nom. (*she*). To avoid possible reference to female characters in the story, all characters in the *them*-conditions were male, so e.g. for (17) and (18) all female names were replaced by male names and the masculine form of *friend* ('vriend') was used. This also avoided two consecutive appearances of 'ze' as in *Laura said that she* ('ze') *them* ('ze') *had found*. The Dutch pronoun in the singular continuation (the *it*-conditions) was the reduced form of the third person singular masculine acc. pronoun 'hem' (*him*): ''m', which (next to referring back to a male person) is used to refer back to a masc./fem. nouns in Dutch (nouns that go with the definite determiner 'de').[8] To avoid reference of ''m' to male characters in the story, all character names were female in the singular continuation condition, and the feminine form of *friend* ('vriendin') was used. In order to make sure participants would not be surprised by the use of the reduced form ''m', which is less frequent in written language than in spoken language, it was included in one of the practice trials.

Both the FocRel condition and the non-FocRel condition came in two versions: one with a plural continuation and one with a singular continuation. Similar to Experiment 8 I also included control conditions with *A and B*, where all occurrences of *or* were replaced by *and*. I only included

[8] In Dutch the equivalent of *it* ('het') is only used to refer back to grammatically neuter nouns, which go with the determiner 'het'.

versions of these conditions with a plural continuation (*them*), as *it* cannot refer back to *A and B*. The total number of conditions therefore was six, exemplified in (21)-(26):

(21) Condition 1: FocRel or them
Karl told a friend (masc.) that Hugo and Oliver went searching for marine animals on the beach yesterday, and that when they returned, their father said that if one of them had found **at least two marine animals** on the beach, that person got to stay up late that night.
The friend (masc.) said: "And, **what had Hugo found** on the beach?"
Karl answered: "Hugo had found a starfish **or** a mussel on the beach."
Hugo said later that he had found **them** in the surf near the lighthouse.

(22) Condition 2: FocRel or it
Marije told a friend (fem.) that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **at least two marine animals** on the beach, that person got to stay up late that night.
The friend (fem.) said: "And, **what had Laura found** on the beach?"
Marije answered: "Laura had found a crab **or** a mussel on the beach."
Laura said later that she had found **it** in the surf near the lighthouse.

(23) Condition 3: non-FocRel or them
Karl told a friend (masc.) that Hugo and Oliver went searching for marine animals on the beach yesterday, and that when they returned, their father said that if one of them had found **a starfish or a mussel** on the beach, that person got to stay up late that night.
The friend (masc.) said: "And, **who of them had found a starfish or a mussel** on the beach?"
Karl answered: "Hugo had found a starfish **or** a mussel on the beach."
Hugo said later that he had found **them** in the surf near the lighthouse.

(24) Condition 4: non-FocRel or it
Marije told a friend (fem.) that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **a starfish or a mussel** on the beach, that person got to stay up late that night.
The friend (fem.) said: "And, **who of them had found a starfish or a mussel** on the beach?"
Marije answered: "Laura had found a crab **or** a mussel on the beach."
Laura said later that she had found **it** in the surf near the lighthouse.

(25) Condition 5: FocRel and them
Karl told a friend (masc.) that Hugo and Oliver went searching for marine animals on the beach yesterday, and that when they returned, their father said that if one of them had found **at least two marine animals** on the beach, that person got to stay up late that night.
The friend (masc.) said: "And, **what had Hugo found** on the beach?"
Karl answered: "Hugo had found a starfish **and** a mussel on the beach."
Hugo said later that he had found **them** in the surf near the lighthouse.

(26) Condition 6: non-FocRel and them
Karl told a friend (masc.) that Hugo and Oliver went searching for marine animals on the beach yesterday, and that when they returned, their father said that if one of them had found **a starfish and a mussel** on the beach, that person got to stay up late that night.
The friend (masc.) said: "And, **who of them had found a starfish and a mussel** on the beach?"
Karl answered: "Hugo had found a starfish **and** a mussel on the beach."
Hugo said later that he had found **them** in the surf near the lighthouse.

4.2 Predictions

Let us first consider the predictions for the *A or B* region. At the point at which *A or B* is read, condition 1 and condition 2 are identical (except for the name changes), as are condition 3 and condition 4. Pooling conditions 1 and 2 (C1+2) and conditions 3 and 4 (C3+4) results in the same design as Experiment 8, with FocRel and connective type (*and/or*) as factors. We again expect an effect of repetition, reflected by a main effect of FocRel: In the non-FocRel conditions the story before the target sentence already contains two appearances of *A or B*, in the conditional and in the question, versus none in the FocRel conditions. Therefore, the critical prediction is again an interaction of FocRel and connective type. The extra processing cost of the SI should result in a bigger difference between the two *or*-conditions (C1+2 vs. C3+4) than between the two *and*-conditions (C5 vs. C6). (Equivalently, the difference between the *or* and *and*-conditions is predicted to be bigger for the FocRel conditions (C1+2 vs. C5) than the non-FocRel conditions (C3+4 vs. C6)).

As for the reading times of *them/it*, let us first consider the four *or*-conditions (C1-C4): There, we expect an interaction of FocRel and pronoun type (*them/it*). Due to the SI we expect *them* to be read slower after C1 (FocRel) than after C3 (non-FocRel) and no such difference is expected for *it* (C2 vs. C4), for which it does not matter whether the SI was calculated.

Accordingly, we expect *them* to be read slower than *it* in the FocRel conditions (C1 vs. C2), but not in the non-FocRel conditions (C3 vs. C4).

As for the four *them*-conditions (C1, C3, C5, C6), we predict an interaction of FocRel and connective type. Due to the SI in C1, *them* should be read slower in the FocRel condition than in the non-FocRel condition after *or* (C1 vs. C3), but this difference should not be observed after *and* (C5 vs. C6), as *them* is never unexpected after *A and B*. Accordingly, we expect *them* to be read slower after *A or B* than after *A and B* in the FocRel conditions (C1 vs. C5), but not in the non-FocRel conditions (C3 vs. C6).[9]

4.3 Critical region and spillover regions

(27) is an example of the segmentation of the items. Slashes indicate segment breaks and double slashes indicate line breaks.

(27) example segmentation Experiment 9
Marije told/to/a/friend/that Laura and Barbara/yesterday/marine animals/ went//
searching/on the beach,/and that/when they returned/their mother/said/ that if/one of them//
at least two marine animals/had found/at/the beach,/that person/later could/stay up.//
The friend/said:/"And,/what/had/Laura/found/at the/beach?"//
Marije answered:/ "Laura/had/a starfish or a mussel/found/at/the/beach."//
Laura/said later/that she/it/had caught/in/the/surf/near the/ lighthouse.//

In all test items, the critical region in the sentence containing the scalar term was the 4th region of the 5th line, the 36th region in total. The regions before the critical region were always the same. The first region (-3) was always the initial storyteller (here: *Marije*) and *answered*, the second region (-2) was always one of the two characters introduced earlier (here: *Laura*), and the third region (-1) was always the auxiliary *had*.[10] In order to reduce variance,

[9] These pairwise predictions are based on the assumptions that:
- Not taking the SI into consideration, the pronoun is read equally fast after an antecedent in the focus as an antecedent in the background (no independent effect of FocRel).
- *It* and *them* are read equally fast (no independent effect of pronoun type).
- *Them* is read equally fast after *A or B* without SI as after *A and B* (no independent effect of connective type).
As we cannot be sure these assumption are true beforehand, the crucial predictions are the interactions.

[10] This is contrary to Experiment 8, where the auxiliary followed the direct object, due to the target sentence being an embedded clause. Here the target sentence is a main clause (as direct speech was used), so the Verb Second rule applies in Dutch and the auxiliary appears in second position.

I applied stricter criteria on the critical region (CR) here than in Experiment 9. The critical region always consisted of 5 words, like *a starfish or a mussel* in (27). It always contained 7 or 8 syllables. Only count nouns with indefinite articles were used. All nouns were 'de' words (common gender) in Dutch. The first spillover region (SO1) was the past participle of the verb, which contained 2-4 syllables (with one exception which contained 5 syllables). The verb form was always identical to the one used in the conditional and the question. I added a PP (e.g. *on the beach*) to the *A or B* sentence to create three more spillover regions. In order for this PP to be natural in the answer, it was also included in the conditional and the question. The second spillover region (SO2) was always a monosyllabic preposition, and the third spillover region was always a (monosyllabic) definite article (Dutch 'de' or 'het'). The last word of the sentence (SO4) was a noun of 1-4 syllables.

The critical region of the pronoun sentence was the 4th region of the 6th line, the 44th region in total, in all test items. The pre-critical regions of the pronoun sentence were also kept constant over items. The first region (-3) always consisted of the subject of the answer (here: *Laura*), the second region (-2) was always *said later*, and the third region (-1) was always *that she/that he* (depending on the condition). The critical region only consisted of the pronoun *it* ('m) or *them* ('ze'). Spillover region 1 (SO1) was always the auxiliary and the past participle of the verb, which was always 2-4 syllables (with one exception, see above) and identical to the form in the conditional and in the question. SO2 was a monosyllabic preposition in all items and SO3 a (monosyllabic) definite or indefinite article ('de', 'het' or 'een'). Finally, SO4 contained a noun of two syllables. The regions following SO4 were not controlled for number of syllables or syntactic category.

4.4 Design

36 story 6-tuples were created as in (21)-(26), and divided over 6 lists in a Latin-square design. The items are given in Appendix 7. Every list therefore contained 6 items per condition. The lists also contained 28 fillers, which were similar to the test items but did not contain *A or B* in the answer.[11] The structure of the fillers was similar to that of the test items. They also started with *X told a friend that Y and Z*, followed by some sort of conditional

[11] Actually, two fillers did contain *A or B* in the answer, but not as the direct object in an episodic sentence (*Emily wanted to adopt a baby from Ivory Coast or Senegal*, *Joop maybe wanted to buy an Audi or a Volkswagen*).

(sometimes containing *A or B*), a question of the friend and an answer by X, followed by a final sentence. Again, typical *who*-questions and typical *what*-questions were used, as well as other types of questions (*how*-questions, yes/no-questions). The final sentence was also frequently uttered by someone else in the story (e.g. the friend who had asked the question).

I created verification statements for 24 out of 64 items (fillers and test items), which were about different parts of the stories. As the experiment was quite long, I decided to include a break halfway through the experiment. For that reason the lists were divided in two. Which part was presented before the break and which after the break was counterbalanced over participants. Within list parts, a newly randomized order of items was determined for every participant by the software.

4.5 Participants and procedure

67 students of Utrecht University, all native speakers of Dutch, participated in the experiment. None of them had any prior knowledge of the topic. None of them had any reading problems. Participants were paid for their participation. The experiment was conducted on computers in sound proof booths, and ran on E-Prime software. The experimental paradigm was the same as Experiment 8 (non-cumulative moving-window paradigm). Participants pressed the space bar with their dominant hand to progress through the text. They were instructed beforehand to read at a normal pace but to read all the material in a segment before moving on. The experiment contained three practice trials, two of which were followed by a verification statement. The procedure for answering the verification statements and proceeding to the next item was the same as in Experiment 8. Halfway through the experiment a pause screen was displayed, and participants were instructed to read magazines for 5 minutes, after which they continued the experiment. The first two items after the break were two additional filler trials, which were the same in all lists. These were included to make sure participants were focused by the time they got to the first test item after the break. Including the break, most participants completed the experiment in around 40 minutes.

4.6 Results *A or B/A and B* sentence

One participant was excluded from the analysis because she scored less than 75% correct on the control statements (71%). The number of participants for the analysis therefore was 66. For easy comparison to Experiment 8, I will

discuss the results of the sentence containing the scalar term in this section (4.6), and the results of the pronoun sentence in 4.7 below.

The same outlier correction was applied as in Experiment 8. Reading times under 200 ms. were removed from the data (0.5% of the data), and reading times that were more than 2.5 SD higher than the grand mean of their region were smoothed to that value (2.4% of the data).

4.6.1 Results *A or B/A and B* region

The smoothing point of the critical region was 2075 ms. 2.1% of the critical region data was smoothed to this value. The average reading times of the *A or B/A and B* region over all participants and all items are given in Table 3 for the six conditions separately.

Table 3: average reading times *A or B/A and B* region Experiment 9

	reading time (ms)
C1 (FocRel or them)	914
C2 (FocRel or it)	896
C3 (non-FocRel or them)	609
C4 (non-FocRel or it)	593
C5 (FocRel and them)	842
C6 (non-FocRel and them)	614

As at the critical region C1 and C2 were identical (except for the name changes), and their reading times did not differ significantly, these were treated as one condition in the analysis. The same held for C3 and C4. The average reading times over all participants and all items of the remaining four conditions are given in Table 4.

Table 4: average reading times *A or B/A and B* region Exp. 9 pooled

	reading time (ms)
C1+2 (FocRel or)	905
C3+4 (non-FocRel or)	601
C5 (FocRel and)	842
C6 (non-FocRel and)	614

A 2x2 ANOVA over subjects (F1) and items (F2) with FocRel and connective type as factors revealed a main effect of FocRel: F1(1,65) =

112.93, p < 0.001, F2(1,35) = 202.86, p < 0.001. This reflects the repetition effect, which was much bigger than in Experiment 8 (differences of 304 ms. and 228 ms. versus 82 ms. and 121 ms. in Experiment 8). No significant main effect of connective type was observed: F1(1,65) = 3.25, p = 0.076, F2(1,35) = 2.48, p = 0.124. Crucially, there was a significant interaction of FocRel and connective type in the subjects analysis: F1(1,65) = 7.67, p = 0.007, which was close to significance in the items analysis: F2(1,35) = 4.00, p = 0.053. This is the interaction we predicted based on a difference in SI-computation between C1+2 and C3+4, which was absent in C5 vs. C6.

Again, paired t-tests (over participants and over items) were calculated for condition pairs that varied only on FocRel or only on connective. I will again report both uncorrected p-values and p-values corrected for 4 comparisons (Bonferroni), all two-tailed. The t-tests revealed that indeed for both *or* and *and*, the focus condition was read significantly slower: *Or* (C1+2 vs. C3+4): t1(65) = 10.607, p<0.001/<0.001, t2(35) = 14.021, p<0.001/<0.001. *And* (C5 vs. C6): t1(65) = 8.129, p<0.001/<0.001, t2(35) = 7.494, p<0.001/<0.001. In line with the observed interaction of FocRel and connective type, the difference between the two *or*-conditions (C1+2 and C3+4), which was 304 ms, is bigger than between the two *and*-conditions (C5 and C6), which was 228 ms. The t-tests show that in the FocRel conditions, *A or B* (C1+2) was read slower than *A and B* (C5): t1(65) = 3.099, p = 0.003/0.012, t2(35) = 2.117, p = 0.041/0.164, while in the non-FocRel conditions there was no significant difference between *A or B* (C3+4) and *A and B* (C6) (both t's <1).

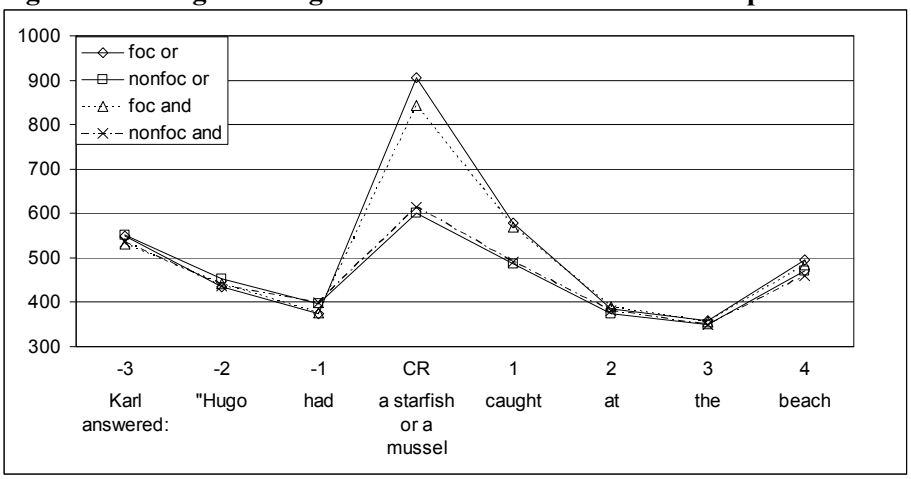
4.6.2 Reading times whole *A or B/A and B* sentence

The average reading times (calculated over all participants and all test items) of the segments in the target sentence are given in Table 5 and Figure 3:

Table 5: average reading times *A or B/A and B* sentence Experiment 9

	-3	-2	-1	CR	SO1	SO2	SO3	+4
C1+2	548	435	373	905	579	384	359	496
C3+4	552	453	397	601	485	374	350	470
C5	531	441	377	843	570	391	357	485
C6	536	438	400	614	491	381	350	459

Figure 3: average reading times *A or B/A and B* sentence Experiment 9



The main effect of FocRel carried over to all four SO-regions. T-tests over participants and items revealed significant differences between C1+2 and C3+4 in the following spillover regions, where * means $p < 0.05$ and ** means $p < 0.013$ (the corrected p-value for 4 comparisons): SO1 (t1** and t2**), SO2 (t1* and t2*), SO3 (t1* and t2**), SO4 (t1** and t2**). The difference between C5 and C6 was significant in SO1 (t1** and t2**) and in SO4 (t1**, t2*). The effect of SI-calculation, the interaction between FocRel and connective type, was not significant in any of the spillover regions.

Let us finally consider the preceding regions. Interestingly, the subject region (-2) showed the same pattern as in Experiment 8, although it was a lot weaker: the non-FocRel *or*-condition (C3+4) was read significantly slower than the FocRel *or*-condition (C1+2) ($t1(65) = -2.632$, $p = 0.011/0.044$, $t2(35) = -2.545$, $p = 0.016/0.064$), and without correction also (marginally) significantly slower than the non-FocRel *and*-condition (C6) ($t1(65) = 2.300$, $p = 0.025/0.100$, $t2(35) = 1.926$, $p = 0.062/0.248$). In section 3.5.2 above I already discussed a possible explanation for this effect. In the -1 region, which consisted of the auxiliary, the pattern 'normalized' as both non-FocRel conditions were read slower than their FocRel counterparts (C3+4 vs. C1+2: $t1(65) = -4.154$, $p<0.001/=0.001$, $t2(35) = -5.082$, $p<0.001/<0.001$, C6 vs. C5: $t1(65) = -3.798$, $p<0.001/=0.002$, $t2(35) = -4.013$, $p<0.001/=0.002$), and a main effect of FocRel was observed: $F1(1,65) = 38.53$, $p < 0.001$, $F2(1,35) = 36.60$, $p < 0.001$. In 3.5.2 above I indicated that this pattern is expected as the subject in the non-focus conditions is actually the information focus. Therefore, it is not surprising the following auxiliary is read slower. The fact that the difference between the non-FocRel *or*-condition (C1+2) and the non-FocRel *and*-condition (C6) of the -2 region

disappeared in the -1 region rules out that it interfered with the effect in the critical region. In the -1 region only a main effect of FocRel was observed, no effect of connective type or an interaction between connective type and FocRel. Therefore, this cannot de responsible for the observed interaction of FocRel and connective type in the critical region.

4.7 Results *them/it* sentence

4.7.1 Results *them/it* region

The smoothing value for the *them/it* region was 651 ms, 2.4% of the reading times were smoothed to that value. The average reading times are given in Table 6.

Table 6: average reading times *them/it* region Experiment 9

	reading time (ms)
C1 (*them* after FocRel or)	373
C2 (*it* after FocRel or)	361
C3 (*them* after non-FocRel or)	358
C4 (*it* after non-FocRel or)	362
C5 (*them* after FocRel and)	356
C6 (*them* after non-FocRel and)	366

2x2 ANOVAs were calculated over subjects and items for the four *or*-conditions (C1-C4) with FocRel and pronoun type as within-subjects and within-items factors, and for the four *them*-conditions (C1,C3,C5,C6) with FocRel and connective type as within-subjects and within-items factors. I will first consider the *or*-conditions. There was a main effect of FocRel: $F1(1,65) = 4.43$, $p = 0.039$, $F2(1,35) = 4.75$, $p = 0.036$, but no main effect of pronoun type: $F1(1,65) = 1.27$, $p = 0.263$, $F2(1,35) = 0.76$, $p = 0.388$. In line with the predictions, there was an interaction of FocRel and pronoun type (marginally significant over items): $F1(1,65) = 4.95$, $p = 0.030$, $F2(1,35) = 5.68$, $p = 0.063$.

In line with the predictions in section 4.2, paired t-tests were conducted on condition pairs that varied in one factor (FocRel, pronoun type, connective), resulting in 7 comparisons. Therefore, I will report uncorrected p-values and p-values corrected for 7 comparisons (Bonferroni correction). Paired t-tests show *them* is read significantly slower after FocRel *or* (C1) than after non-FocRel *or* (C3): $t1(65) = 2.805$, $p = 0.007/0.049$, $t2(35) = 2.954$, $p = 0.006/0.042$, while this difference is not observed for *it* (C2-C4:

both t's < 1). Accordingly, after FocRel *or*, *them* (C1) was read slower than *it* (C2), although this difference was only significant in the over participants analysis without correction: t1(65) = 2.386, p = 0.020/0.140, t2(35) =1.839, p = 0.074/0.518. No such difference was observed for the non-FocRel conditions (C3-C4: both t's < 1). We can conclude that in line with the predictions, the delay was only observed in the FocRel condition, and only for *them*, not for *it*.

The 2x2 ANOVAs on the four *them*-conditions (C1, C3, C5, C6) reveal no main effect of FocRel (both F's < 1), nor of connective type: F1(1,65) = 1.31, p = 0.257, F2(1,35) = 1.42, p = 0.241. However, the interaction of FocRel and connective type that we predicted was highly significant: F1(1,65) = 13.30, p = 0.001, F2(1,35) = 9.37, p = 0.004.

In the t-tests above we already saw that *them* was read slower after FocRel *or* (C1) than non-FocRel *or* (C3). As is immediately clear from Table 6, this was not the case for the two *and*-conditions (C5-C6). The difference between the two actually went the other way but was not significant: t1(65) = -1.804, p = 0.076/0.532, t2(35) = -1.593, p = 0.120/0.840). The t-tests revealed that accordingly, *them* was read slower after FocRel *or* (C1) than FocRel *and* (C5) (t2 marginally significant with correction): t1(65) = 3.190, p = 0.002/0.014, t2(35) = 2.784, p = 0.009/0.063, while no such difference is observed for the non-FocRel conditions. Actually, *them* is read faster after non-FocRel *or* than after non-FocRel *and*, but this difference is not significant (C3-C6: t1(65) = -1.426, p = 0.159/1.000, t2(35) = -1.456, p = 0.154/1.000). We conclude that in line with the predictions, the delay was only observed in the FocRel condition, and only after *or*, not after *and*.

4.7.2 Reading times whole *them/it* sentence

Table 7 shows the reading times of the whole *them/it* sentence. To avoid overcrowding one graph, Figure 4 presents the reading times in the four *or*-conditions (C1-C4) and Figure 5 the four *them*-conditions (C1,C3,C5,C6), in line with the analysis.

Table 7: average reading times *them/it* sentence Experiment 9

	-3	-2	-1	CR	SO1	SO2	SO3	SO4	5	6
C1	458	388	368	373	414	387	367	394	450	653
C2	449	385	362	361	389	372	362	385	443	623
C3	462	377	356	358	389	377	355	381	445	642
C4	452	384	360	362	385	370	355	382	441	625
C5	446	379	363	356	403	371	351	381	423	642
C6	455	391	371	366	395	371	353	378	421	629

Figure 4: average reading times *them/it* sentence Exp. 9 (*or*-conditions)

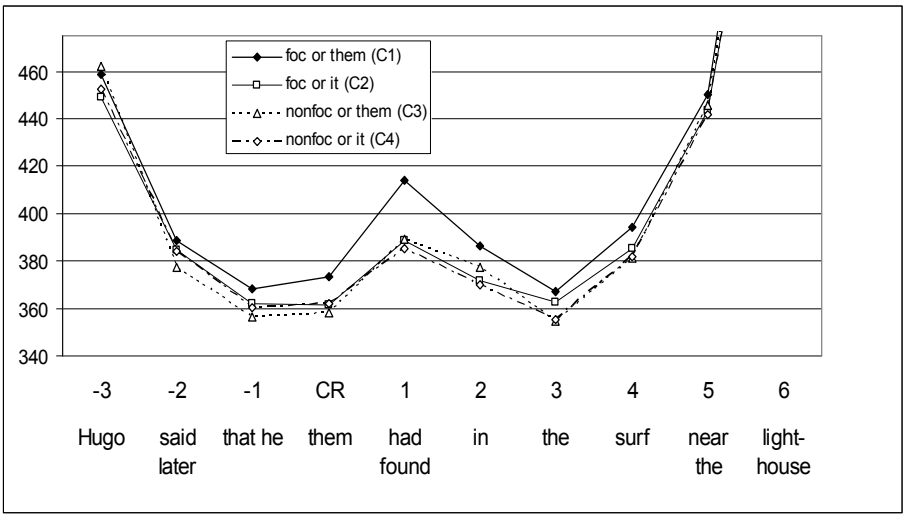
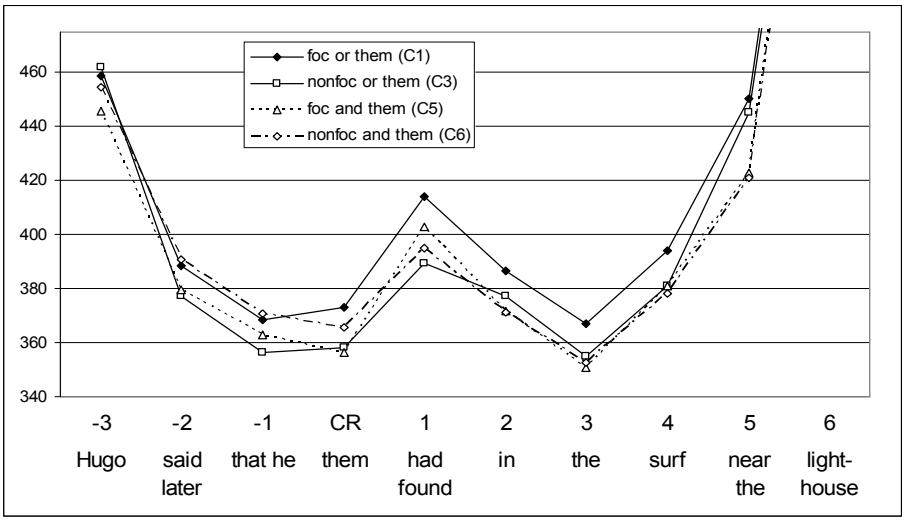


Figure 5: average reading times *them/it* sentence Exp. 9 (*them*-cond.)



Let us first consider the spillover regions. The delay on C1 spilled over to the four spillover regions, indicating participants had problems finding an antecedent for *them* after *A or B* in the FocRel condition, but not in any of the other conditions. C1 was read significantly slower than C3 in the following regions, where * means $p < 0.05$ and ** means $p < 0.007$ (the corrected p-value for 7 comparisons): SO1(t1** and t2**), SO2 (t2*), SO3 (t1* and t2**), SO4 (t1* and t2*). C2 was not significantly different from C4 in any spillover region. C1 was read significantly slower than C2 in: SO1 (t1** and t2*) and SO2 (t1** and t2**), while reading times in C3 were not significantly different from C4 in any of the spillover regions. This indicates that also in the spillover regions, we only observed an effect in the FocRel condition, and only for *them*, not for *it*.

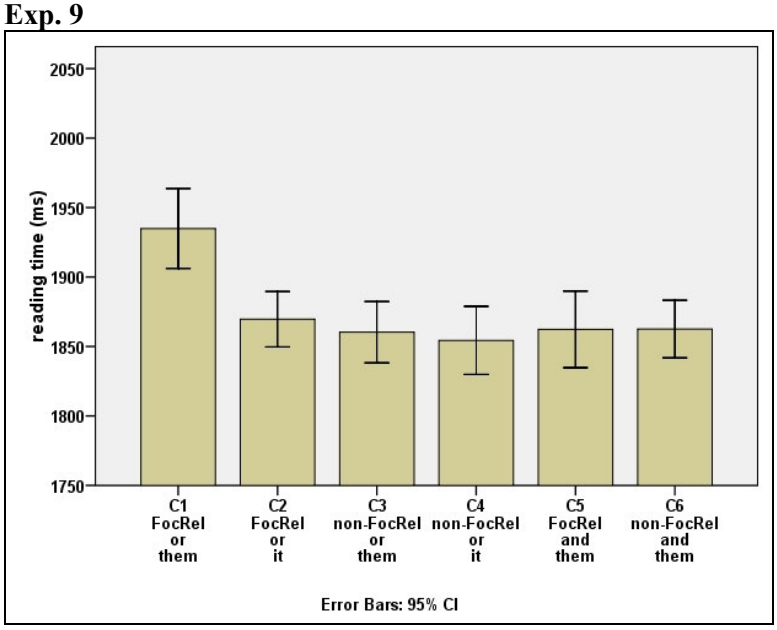
C5 and C6 also did not differ significantly in any region, nor did C3 and C6. C1 however returned slower reading times than C5 in SO2 (t1* and t2*), SO3 (t1** and t2**) and SO4 (t2*). So also in the spillover regions the delay was only present in the FocRel condition, and only after *or*, not after *and*.

If we add up the reading times of the critical region and the spillover regions, as is done in Table 8, a clear pattern emerges which is represented in Figure 6. C1 was slower than all other conditions, while the other conditions returned very similar reading times to each other. This is in line with our prediction that the SI is only calculated in C1, leading to problems in finding an antecedent for the plural pronoun *them*.

Table 8: average cumulative reading times *them/it* and spillover regions Exp. 9

	total reading time CR+SO1-4
C1 (*them* after FocRel or)	1935
C2 (*it* after FocRel or)	1870
C3 (*them* after non-FocRel or)	1860
C4 (*it* after non-FocRel or)	1854
C5 (*them* after FocRel and)	1862
C6 (*them* after non-FocRel and)	1863

Figure 6: average cumulative reading times *them/it* and spillover regions Exp. 9



In the regions after the spillover regions one interesting difference was observed. In the +5 region, which was not controlled for syntactic category or syllable length, the two *and*-conditions were read 20-30 milliseconds faster than the four *or*-conditions. This might reflect processes of discourse integration. It might be easier to integrate discourse referents that were (part of) *A and B* than *A or B*, due to the indeterminacy of *A or B*. I will not discuss this observation any further, as I think it is not related to the current topic.

Two differences were observed in the pre-critical regions. The -2 region (*said later*), was read significantly slower after non-FocRel *and* (C6) than after non-FocRel *or* (C3), but only without correction (t1(65) = -2.132, p = 0.037/0.259, t2(35) = -2.038, p = 0.049/0.343). In the -1 region this difference was also significant without correction only (t1: p = 0.024/0.168 and t2: p = 0.042/0.294). It is unclear whether this difference is meaningful, but the only difference between the two conditions at that point is the use of *and* vs. *or* in the conditional and the question. However, at that point C3 and C4 are still identical (except for the name changes) but contrary to C3, the difference between C4 and C6 does not reach significance in the -2 or the -1 region. As the difference is possibly due to chance and as I do not think it is relevant for our hypothesis, I will not look into it further.

The second difference in the pre-critical regions was that the -1 region (*that he/that she*) was read significantly slower in C1 than C3. However, this difference was also only significant without correction: (t(65) = 2.416, p = 0.019/0.133, t(35) = 2.296, p = 0.028/0.196). Even though this difference was small, it is potentially worrying, as the difference between these conditions in the critical region was one of the important observations supporting our hypothesis. However, for our conclusions it is important that the difference between C1 and C3 was much more pronounced at the critical region. It is not until the critical region that the difference is significant after correction for multiple comparisons. Therefore, the difference in the -1 region might simply be due to chance. In SO1 the absolute difference further increased, and there the difference was also highly significant (t1: p = 0.004/0.028 and t2: p = 0.001/0.007). That region contained the auxiliary and the verb so at that point participants could be sure *them/it* referred back to *A or B/A and B*. This suggests the effect was indeed due to problems with finding an antecedent for *them* after an SI was calculated. A final argument for why any effect in the -1 region could not have been responsible for the observed effect at the critical region is that the other differences that were observed in the critical region (C1 slower than C2, C1 slower than C5) were not observed in the -1 region (all p's > 0.200 (uncorrected)).

5. General discussion Experiments 8 and 9

I present the general discussion in two parts. In 5.1 I address the implications of the results for the hypothesis that SIs are focus-sensitive. In 5.2 I discuss how the data contribute to the discussion about the processing cost of SIs.

5.1 Implications of the results about the focus-sensitivity of SIs

In this section I will consider the implications of the results of Experiments 8 and 9 for the hypothesis of focus-sensitivity of SIs. I address the question whether the results support the view that SIs are costly in the next section.

In Experiment 9 the reading times on the region containing the scalar term *or* showed a significant interaction of FocRel and connective type. Only in the FocRel-conditions we observed slower reading times for *A or B* than *A and B*. One difference between *and* and *or* is that in upward entailing episodic sentences *or* can trigger an SI while *and* cannot. Therefore, the extra processing load observed in the FocRel *or*-condition suggests that more SIs were calculated in this condition than in the non-FocRel *or*-condition, which did not differ in reading times from the non-FocRel *and*-condition. This supports the hypothesis that the cumulative property of focus and alternative relevance increases SI-calculation.

In the sentence following the scalar sentence in Experiment 9, the pronoun *them* and the spillover regions were read slower after the FocRel *or*-condition than the non-FocRel *or*-condition. This difference was not observed for the control conditions with *it*, nor for the control conditions with *and*. This indicates that the sentence containing *or* was more frequently interpreted with an SI in the FocRel condition than in the non-FocRel condition. This again supports the hypothesis that focus and alternative relevance together increase SI-calculation.

In Experiment 8 no delay on the scalar region was observed in the focus condition compared to the non-focus condition. It is tempting to conclude from the difference between Experiments 8 and 9 that focus alone is not enough to trigger SIs, and that alternative relevance is also needed. However, this conclusion crucially depends on whether the delay on the scalar region that was observed in Experiment 9 was indeed due to SI-calculation. The fact that it co-occurred with a delay on the plural pronoun suggests it was indeed due to SI-calculation, so the fact that no delay on the scalar region was observed in Experiment 8 could indicate that there was no difference in SI-calculation between the conditions there.

However, we have to be careful in drawing conclusions about the meaning of the delay on the scalar region in Experiment 9. The scalar region was also the part of the sentence where reasoning about the outcome of the conditional could begin. So next to possible SI-calculation, reading times on this region might also reflect inferential processes about whether or not the requirement in the conditional was satisfied. In this light it is crucial to realize that the FocRel *or*-condition, which was the only condition in which a delay was observed, was the only condition in which it was not clear from the lexical meaning of *A or B* (inclusive-*or*) whether the requirement was

satisfied. The requirement was (e.g.) to find at least two animals, so on an inclusive reading of *or* it was unclear whether the requirement was satisfied. Inclusive-*or* allows for the situation in which both were found and the situations in which one of the two was found. Only after SI-calculation it could be concluded that it was not satisfied. In the non-FocRel *or*-condition, this conclusion could be drawn immediately from the lexical meaning: The requirement was (e.g.) finding a crab and/or a mussel, so an inclusive reading of *or* in the answer was enough to determine that it was satisfied. Also in both *and*-conditions, the lexical meaning of *and* was enough to determine that the requirement had been satisfied (in both conditions). Furthermore, the FocRel *or*-condition was the only condition in which the requirement of the conditional was (potentially) not satisfied, as calculating the SI results in the requirement not being satisfied. Therefore, we cannot rule out that the delay is due to differences in inferential processes about the outcome of the conditional between the FocRel *or*-condition and the other conditions, instead of due to SI-calculation itself. I do not claim this is a better explanation for the delay than SI-calculation, but based on the current experiment it cannot be ruled out. This would explain why no effect was found in Experiment 8. If the delay on the scalar region in Experiment 9 is due to inferences about the outcome of the conditional, it makes sense that no delay was observed in Experiment 8, as the stories did not contain conditionals there. Therefore, we cannot conclude from the difference between Experiments 8 and 9 that focus alone is not enough to trigger SIs.

There is another reason why the difference between Experiments 8 and 9 does not point to alternative relevance being more important than focus. Recall that the initial reason to involve alternative relevance (in chapter 4) was that it was an alternative explanation for the data of the TVJT-experiments. There, the effect could have also been due to the question in the focus condition (*What did Katja find?*) triggering alternative relevance, while the question in the non-focus condition (*Who found a crab or a starfish?*) did not. On the view that alternative relevance is the crucial property and not focus, this has to be the explanation of the effect of the TVJT-experiments. As no conditionals were present in e.g. (TVJT) Experiment 2, the effect has to be due to alternative relevance triggered by the question. But if that was the case, why did we not observe an effect in Experiment 8, in which the same questions were used so the same difference in alternative relevance was present? We would have to resort to the claim that because the *A and B* situation was taken out in Experiment 8, there was no incentive to calculate the SI anymore, despite the alternative relevance difference triggered by the different questions. But then we are back where we started and the difference between Experiments 8 and 9 tells us nothing

about which property (focus or alternative relevance) is the crucial one anymore.

A way to assess the hypothesis that focus alone is not enough in a follow-up experiment is by using the measure of the plural pronoun continuation on a condition in which alternative relevance and focus are contrasted, as in condition 1 of the PWJT-experiments. The critical item could look like (28), which is the FocRel *or*-condition of Experiment 9 with *at least two animals* replaced by *a starfish or a mussel* (given in boldface):

(28) [-alternative relevance, +focus] condition
Karl told a friend (masc.) that Hugo and Oliver went searching for marine animals on the beach yesterday, and that when they returned, their father said that if one of them had found **a starfish or a mussel** on the beach, that person got to stay up late that night.
The friend (masc.) said: "Oh, and what had Hugo found on the beach?"
Karl answered: "Hugo had found a starfish or a mussel on the beach."
Hugo said later that he had found them in the surf near the lighthouse.

If we find the delay on *them* in this condition too, focus alone is enough to trigger an SI, as with this conditional the context has no alternative relevance. If the delay is not observed here, it seems alternative relevance is needed. I leave this as a suggestion for further research.

Notice that the possibility that the delay on the scalar region is due to something else does not harm the conclusion of Experiment 9 that focus and alternative relevance together increase SIs. This also followed from the delay on the plural pronoun *them*, which is independent of the issue what caused the delay on the scalar region. I will now consider the implications of the results for the processing cost of SIs.

5.2 Implications of the results about the processing cost of SIs

The fact that in the same condition in which a delay was observed on the plural pronoun, a delay on the scalar region was observed, suggests that SI-calculation is costly. This is in line with the contextualist view, who claim SIs are costly inferences that are only calculated when they are licensed by the context. In the previous section I already indicated we have to be careful in attributing the delay on the scalar region to calculation of SIs, as alternative explanations are possible. However, in this section I will present an argument that the results nevertheless support the contextualist view that SIs are costly over the defaultist view that SIs are cheap, based on a

comparison of the results of the TVJT-experiments and the current experiments.

If the delay found on the scalar region in Experiment 9 is the result of calculation of an SI, this is in favor of the contextualist account and goes against the defaultist account. If the delay is the result of something else, e.g. inferences about the outcome of the conditional, the default account is still possible. However, if the defaultist view is right, the question arises why in neither of the two experiments we observed a delay due to the cancellation of the SI in the non-focus *or*-condition compared to the non-focus *and*-condition. This will be a crucial argument for why the data support the contextualist view.

For Experiment 8, the defaultists could make a similar argument about the absence of an SI-cancellation effect that I made about the absence of an SI-calculation effect. I proposed that there was no reason to calculate the SI in the focus *or*-condition, as nothing hinged on the SI, contrary to the TVJT-experiments, in which the SI was important for checking the target sentence with the *A and B* situation. Similarly, the defaultists could propose that there was no cancellation in the non-focus *or*-condition in Experiment 8 because nothing hinged on the SI. However, this conclusion is too hasty. Remember that on a defaultist account calculation is automatic, so we have to look at the *forces that drive cancellation*.

Let us compare Experiment 8 to (TVJT) Experiment 2, as the only difference between the stories of the two experiments is that the actual *A and B*-situation was taken out in Experiment 8. In Experiment 2, participants had an incentive to calculate the SI because the target sentence with *or* had to be checked with the *A and B*-situation. According to the defaultists, the SI of *or* is calculated by default. However, despite the fact that the SI is relevant, the fact that *or* is non-focus triggers cancellation in a fair number of cases (59%, as the SI-rate in the non-focus condition was 41%). So the cancellation is driven by non-focus, against the relevance of the SI due to the task. In Experiment 8 the participant no longer had to check the target sentence with the *A and B* situation, decreasing the relevance of the SI compared to Experiment 2. As the SI is less relevant here than in Experiment 2, the force opposing the cancellation-by-non-focus is smaller, so we would expect the effect of non-focus triggering cancellation to be even bigger here. So while the SI not being relevant in Experiment 8 was a possible explanation for the contextualists for why there was no delay in the focus *or*-condition, it is not a possible explanation for the defaultists for the absence of a delay in the non-focus *or*-condition.

Experiment 9 also showed no delay in the non-focus *or*-condition compared to the non-focus *and*-condition, while this was expected on the defaultist view (due to cancellation of the SI in the non-focus *or*-condition).

It did show that a plural pronoun was read faster after non-focus *or* than focus *or*, indicating that the SI was absent in the non-focus *or* case, which on the defaultist-view means it was cancelled. This cancellation was however not reflected in a delay in reading times of the non-focus *or*-condition compared to the non-focus *and*-condition in any region from *A or B* till *them*. We can conclude that the results of Experiments 8 and 9 favor the contextualist 'only calculate costly SIs when necessary'-view over the defaultist 'SIs are cheap and automatic but can be cancelled'-view.

In the next chapter I extend the investigation by testing the hypothesis of focus-sensitivity of SIs on another scalar term (*most*) in another structural position. Despite its shortcomings discussed earlier, I will use the TVJT for easy comparison to the data of *or* in Experiments 1 and 2.

CHAPTER 7: BROADENING THE SCOPE: THE *MOST*-EXPERIMENTS

1. Introduction

In Chapters 3, 4 and 6 I tested the hypothesis that SIs are focus-sensitive by considering answers to wh-questions in which the scalar term *or* appeared in direct object position. In this chapter I broaden the scope of the investigation by presenting two off-line experiments in which two factors were changed relative to the experiments presented so far. The first change was that in the experiments another scalar term than *or* was tested: *most*. The second change was that the experiments tested a scalar term in subject position instead of object position. Additionally, I included a new environment in which the focus-sensitivity of SIs was assessed, answers to yes/no-questions. Finally, I added one condition to explore the interpretation of scalar terms in yes/no-questions themselves, independent of the focus hypothesis. The goal of these experiments was twofold. The first goal was finding additional evidence for the hypothesis of focus-sensitivity of SIs by replicating the results with another scalar term in another scalar position, and by testing whether the focus predictions match the interpretation of scalar terms in answers to yes/no-questions. The second goal was exploring the interpretation of scalar terms in questions themselves.

The written TVJT-experiments (Experiments 1 and 2) of chapter 3 were the reference point for the experiments presented in this chapter, Experiments 10 and 11. These experiments were also conducted with a written TVJT. In chapter 3 I pointed to the weaknesses of the TVJT to assess SIs. However, as the TVJT experiments provided clearer results than the experiments in which the other off-line paradigm was used (the PWJT), it is the best point of comparison. Therefore, I used the TVJT despite its methodological shortcomings. However, in Experiment 11 I used a slightly different version of the TVJT, asking participants to judge the target sentence 'right' or 'wrong' instead of 'true' or 'false'. Hence, the methodological goal of finding the right paradigm to assess SIs is not completely ignored in this chapter.

In section 2.1 I show how the two changes (another scalar term in another scalar position) mentioned above were implemented in the experiments In section 2.2 I introduce the new environment in which the focus-hypothesis will be tested (yes/no-questions) and in section 2.3 I discuss the explorative condition which was added (interpretation of scalar terms in yes/no-questions themselves). Experiment 10 will be presented in section 3, followed by Experiment 11 in section 4. In section 5 I provide the

general discussion in three parts (5.1-5.3), corresponding to the three issues introduced in sections 2.1-2.3. I end this chapter by presenting the conclusions in section 6.

2. The two changes, the new condition and the explorative condition

2.1 Another scalar term in another structural position

In order to test the predictions of the focus-sensitive view of SIs with another scalar term, we have to find a suitable alternative to *or*. The obvious candidate for the job is the other hallmark example of a scalar term: English *some*. Sentences like (1) give rise to the SI that not all students (in the relevant domain) drank beer.

(1) Some students drank beer.

However, in Dutch there are two equivalents of *some*: 'sommige' and 'enkele'. Both are not directly comparable to *some*. 'Sommige' is claimed to have a specific flavor: the members of the set have to be related to each other (De Hoop & Kas 1989), in that sense it is similar to English *certain*. 'Enkele' on the other hand is preferably used for small amounts, like English *a few*. Banga et al. (2009) showed that *sommige* and *enkele* give rise to different SI-rates. In order to avoid these problems, I decided to use another quantifier: *most* ('meeste' in Dutch).

Like *some*, *most* is on an entailment scale with *all*, so (2) can also trigger the SI that not all students drank beer (see Horn 1972).

(2) Most students drank beer.

There has been some discussion in the literature on whether the upper-bound of *most* (*not all*) is an SI or whether it is encoded in the semantics. Ariel (2004) claims *most* means something like 51%-99%, so with the upper bound as part of the semantic meaning. However, Horn (2006) and Papafragou & Schwarz (2006) provide arguments against this view and argue in favor of a scalar meaning. I will follow them in this chapter and assume *most* is a genuine scalar term.

I also changed the structural position in which the scalar term appeared. I decided to use sentences like (2) above, in which *most* is part of the subject. This also facilitated a test of the hypothesis that a default preference for a sentence-final focus might have affected the results of the experiments so far (see chapter 3 section 5.3). If this preference increased the SI-rates in the

non-focus conditions, this experiment should return lower SI-rates in this condition.

2.2 The new condition for testing the focus predictions: answers to yes/no-questions

So far, I have only tested the predictions of focus-sensitivity of SIs on answers to wh-questions. In Experiments 10 and 11, I included answers to yes/no-questions as well. The sentence in (2) above can be given as an affirmative sentential answer to a yes/no-question, as in (3):

(3) Q: Did most students drink beer?
A: Most students drank beer.

The answer in (3) contains no new information, all the parts of the sentence were already present in the question. The only new information that is communicated is an implicit 'yes'. Therefore, we could claim the sentence is as a whole focus-less. I will take this as a working assumption, but I will return to it in the discussion in section 5.2. On the assumption that (3A) is as a whole background, the focus-sensitive accounts predict that an affirmative sentential answer to a yes/no-question should return the same (low) SI-rate as when the same sentence is an answer to a wh-question in which the scalar term is part of the background.

2.3 The explorative condition: scalars in questions themselves

The condition I wanted to explore is the interpretation of scalar terms in questions. It is generally claimed in the literature that SIs do not arise in questions. For instance, Noveck et al. (2002) tested propositional reasoning problems like (i):

(i) premise 1: If there is a P then there is a Q and an R.
premise 2: There is a P.
question: Is there a Q or an R?

Only 20% of the 20 participants answered 'no' to the question, indicating that the majority had interpreted *or* in the question inclusively, so without

SI.[1] Contrastively, in another experiment the same premises were used but the conclusion was the declarative sentence *There is a Q or an R*, and 75% of the participants rejected it (presumably based on the SI). Noveck et al. concluded that 'This indicates that the question form serves as a linguistic context in which inclusive interpretations are acceptable' (p. 319). However, in Noveck et al.'s experiment the *and*-situation is the consequent of an abstract conditional. I wanted to investigate how a scalar term is interpreted in a question relative to a (story describing a) real world situation which could be described by the stronger scalar term.

While for wh-questions it is impossible to check the interpretation of a scalar term in the question itself using the TVJT, for yes/no-questions a slight modification of the task makes this possible. Similar to Noveck et al. (2002), instead of having participants judge the answer true or false, I made them answer the question themselves, as in (4).

(4) Q: Did most students drink beer? yes / no

After a story in which all students drank beer, a 'no' answer indicates the participant interpreted *most* as *most but not all*, so with an SI. I now turn to the experiments.

3. Experiment 10: broadening the scope with *most*

3.1 Setup and items

As in (TVJT) Experiment 2, the story was kept constant between conditions, so the effect of focus could be isolated. Example items of the wh-conditions are given in (5) and (6):

[1] In chapter 3 I report data from another experiment of Noveck et al. (2002), in which participants were presented with both questions and declaratives. I will come back to this experiment in the discussion.

(5) Condition 1 (wh focus)
Five people were present at the get-together of the Celtic language studies program at the University. Several drinks were available.
Sander is a student. He drank beer.
Tom is also a student. He also drank beer.
Eric is a professor. He drank wine.
Martin is a professor too. He drank apple juice.
Frans is a student. He drank beer.

A: "How many of the students drank beer?"
B: "Most students drank beer."

(6) Condition 2 (wh non-focus)
Five people were present at the get-together of the Celtic language studies program at the University. Several drinks were available.
Sander is a student. He drank beer.
Tom is also a student. He also drank beer.
Eric is a professor. He drank wine.
Martin is a professor too. He drank apple juice.
Frans is a student. He drank beer.

A: "What did most students drink?"
B: "Most students drank beer.

As I wanted to keep the story constant over conditions, I had to introduce both a contrast set for the subject (several people) as for the object (several drinks). Also, in order for the stories to not become overly simple and boring, I introduced a contrast set for the restrictor of *most*: the story also contains some professors. Consider how obvious the critical inference would have become if this set would have been left out:

(7) Three students were present at the get-together of the Celtic language studies program at the University. Several drinks were available.
Sander drank beer.
Tom drank beer.
Frans drank beer.

A: "How many of the students drank beer?"
B: "Most students drank beer.

Contrary to the previous experiments the question in (5) does not question the whole subject constituent, as in (8):

(8) A: "Who drank beer?"
B: "Most students drank beer."

This choice was made because the *who*-question might introduce confusion due to a 'partially true' answer on the SI-reading: If the dialogue in (8) is combined with the story in (5), it is true in the sense that it provides the right information that it was the *students* that drank beer, not the professors. However, it is not true that *most but not all* of them drank beer. So on an SI-reading, the sentence is partly true and partly false. To avoid 'true' answers while the SI was calculated, I chose to use the *How many* question in (5) instead. As a result, the focus condition is slightly different from the experiments with *or*, as here only the scalar term itself is the focus of the sentence, instead of the whole constituent. However, this does not make a difference for the predictions of the accounts of Van Rooij and Van Kuppevelt, which predict the SI to arise as *most* is in the focus part of the sentence.

However, in section 7.2 of chapter 3 I claimed the experimental items of Schwarz et. al (in progress), in which only the scalar term *or* was stressed (e.g. *Mary will invite Fred OR Sam to the barbecue*), were not comparable to my items in which the whole constituent *A or B* was information focus. I claimed stress on *or* itself as in Schwarz et al.'s items could only trigger a contrastive focus, as the sentence can only be a correction to the same sentence with *and*. This is however not the case with *most*. Even though the answer in (5) would also be pronounced with stress on the scalar term *most* itself, the fact that it is an answer to the question posed in (5) indicates that it is a regular information focus.

In all stories, 5-7 people were introduced, either by name as in (5) and (6), numbered (*The first student...The second student*), or paired (*John... His wife...*). The people in the story were always divided into two groups (in this example students and professors). The explicit question was always about the biggest group (in this example the students), which always consisted of 3-5 people.

The new condition which tested the interpretation of the scalar term in an answer to a yes/no-question, contained the same story as (5) and (6), but was followed by a dialogue of the following type:

(9) <u>Condition 3 (yes/no sentential answer)</u>
A: "Did most students drink beer?"
B: "Most students drank beer."

I will call this condition the *yes/no sentential answer condition*. Notice that the target sentence was the same as in the wh-conditions. As I assume the whole sentence is focus-less, no SIs were predicted.

In the second new condition, which was included to explore the interpretation of *most* in the yes/no-question itself, the same story-type as in the other conditions (see (5) and (6)) was used. The story was followed by the same question as in the yes/no sentential answer condition (see (9)). However, contrary to (9) the answer was left out, and participants were asked to answer the question themselves by choosing 'yes' or 'no', see (10). I will call this condition the *yes/no polar condition*.

(10) Condition 4 (yes/no polar)
A: "Did most students drink beer?" yes / no

3.2 Design

8 stories were created with each 4 dialogues corresponding to the 4 conditions (see Appendix 8). These were divided over 4 lists, so every list contained 2 items per condition. The 8 test items per list were interspersed with 8 fillers, so every list consisted of 16 stories. 6 fillers also contained a target sentence with *most*, and two fillers contained a target sentence with *the minority of*. In the fillers with *most*, the target sentence was either true irrespective of SI (e.g. 3 out of 4), or false irrespective of SI (e.g. 2 out of 5). All the question/answer-types (wh-subject, wh-object, yes/no-sentential, yes/no-polar) were represented evenly in the fillers. The 4 lists each had a different order of items and fillers, with the following restrictions: there were never two consecutive test items of the same condition, never more than two consecutive test items, never more than two consecutive items with the same question-type, and the first test item was of a different condition for each of the four lists.

3.3 Participants and procedure

35 participants completed the web-based questionnaire (WWSTIM, Veenker 2000) on their own computers. All were adult native speakers of Dutch, with no prior knowledge of the topic. Most of them were students or had a university degree. They were instructed that they would read 16 stories, followed by a dialogue between speaker A, who did not know what happened in the story, and speaker B, who did know what happened. Their task was to judge whether speaker B's answer was true, by clicking the

'true' button or the 'false' button, while the story and the dialogue remained on the screen. They were instructed that sometimes they had to answer a question of speaker A themselves, by clicking 'yes' or 'no'. Every item contained a text box under the 'true' and 'false' buttons, in which participants could type comments. Participants were that there were no 'right' or 'wrong' answers and they should follow their intuition. Most participants completed the experiment within 15 minutes, with an average of 11 minutes.

3.4 Results

A number of participants was excluded because of an unavoidable problem with the test items of condition 1, repeated here for convenience:

(5) Condition 1 (wh focus)
Five people were present at the get-together of the Celtic language studies program at the University. Several drinks were available.
Sander is a student. He drank beer.
Tom is also a student. He also drank beer.
Eric is a professor. He drank wine.
Martin is a professor too. He drank apple juice.
Frans is a student. He drank beer.

A: "How many of the students drank beer?"
B: "Most students drank beer."

Some participants seemed to have judged this answer 'false' not because of an SI, but because they felt the speaker should have provided the exact number of students that drank beer (in this case 'three') as an answer, instead of a proportion. This problem could however be controlled for by looking at the fillers and the participants' comments. If participants answered 'false' to items like (5) because no exact number was provided, they should have also done so for the fillers in which *most* was true irrespective of SI (e.g. a situation with 4 out of 5). Therefore, participants were excluded if they judged one or more of these true *most* fillers with a *How many* question false, or if they commented in the textbox that an exact number was required (or both). Five participants were excluded based on these criteria. Three participants commented about an exact number in the textbox but answered 'true' to the true *most* fillers with a *How many* question, so they were not excluded. No participants were excluded based on error rates on the fillers,

as all of the participants' accuracy rates on the fillers was at least 75%. After exclusion the judgments of the remaining 30 participants were analyzed.

The percentages of 'false' answers, indicating the calculation of the SI, are given in Table 1:

Table 1: SI-rates Experiment 10

Condition 1 (wh focus)	Condition 2 (wh non-focus)	Condition 3 (yes/no sent.)	Condition 4 (yes/no polar)
52%	42%	42%	40%

Let us first focus on the two wh-conditions. The difference between the SI-rates of the two wh-conditions is small but significant over participants: Wilcoxon signed rank test gives $z = -2.12$, $p = 0.017$ (one-tailed).[2] Therefore, the effect of focus-sensitivity of SIs we observed for *or* in direct object position is replicated for *most* in subject position.

The distribution of participants over the different possible types of behavior on conditions 1 and 2 shows a clear pattern given in Table 2, where 'more SIs' means at least one 'false' answer more, 'SIs in both conditions' means 2/2 'false' answers in both conditions, and 'SIs in neither' means 0/2 'false' answers in both conditions.

Table 2: Distribution of participants Experiment 10, wh-conditions

more SIs in foc	SIs in both cond	SIs in neither	more SIs in n-foc
23% (7)	27% (8)	47% (14)	3% (1)

The percentage of participants that distinguished between the two conditions in the predicted direction went down from around 50% in the TVJT-experiments with *or* to 23% here. However, the opposite pattern (more SIs in the non-focus condition) is also hardly observed here, which supports the focus-sensitive view. However, especially the group that did not calculate SIs in either condition increased compared to the experiments with *or*.

As predicted, the yes/no sentential answer condition (C3) patterned with the wh non-focus condition (C2), rather than with the wh focus condition (C1). The difference between C3 and C1 was significant: Wilcoxon signed rank test (over participants) $z = -1.86$, $p = 0.032$ (one-tailed).

[2] The difference was not significant over items: $z = -1.025$, $p = 0.153$ (one-tailed). There was more variance in the items-scores in this experiment than in Experiments 1 and 2. This is probably (partly) due to the fact that this experiment contained more conditions and therefore fewer data points per item-condition pair (e.g. 8,75 in the current experiment versus 18,5 in Experiment 1 and 23 in Experiment 2). Therefore, I will only report analysis over participants in this chapter. The interested reader is referred to Appendix 8 for the rates per item.

The polar condition (C4) also returned fewer SIs than the wh focus condition, in line with the view that questions are SI-blocking. However, it returned nearly as many SIs as C3, suggesting that SIs are equally likely to arise in sentential answers to yes/no-questions as in yes/no-questions themselves.

3.5 Discussion

Even though the differences are small, the results suggest that also for another scalar term in another structural position the SI is sensitive to focus, supporting the focus-sensitive views of SIs. However, the difference between the wh focus and the wh non-focus condition is smaller than in the experiments with *or*. I will provide a number of explanations for this in section 5.1 below. The results also show that when the sentence containing the scalar term is an answer to a yes/no-question and as a result focus-less, fewer SIs are observed than if the scalar term is in the focus part of an answer to a wh-question. This is also in line with the predictions of the focus-views. Finally, the results suggest SIs are just as likely to arise in yes/no-questions themselves as in answers to yes/no-questions, but I will return to this in section 5.2 below.

First there is another factor that I would like to consider, which might be responsible for the difference between the conditions in both Experiments 1-3 and Experiment 10 being smaller than predicted by the theory. That is the issue what exactly the participants were asked to judge the target sentence on. In the Experiments 1-3 and Experiment 10, participants were explicitly asked to judge the sentence on *truth*, and ignore strangeness or ugliness of the sentence. The disadvantage of asking participants to judge the sentence on truth alone, is that they might be inclined to consider the target sentence in isolation, ignoring for instance the question it is an answer to. As the critical manipulation in these experiments is the question, the effect would dramatically decrease due to a strategy like this.

To control for this possibility, I conducted a control experiment for Experiment 10 in which participants were asked to judge whether the sentence was 'right' (Dutch: 'goed') or 'wrong' (Dutch: 'fout').[3] [4] The idea

[3] When the TVJT is used in acquisition research children are also usually asked whether the sentence is 'right' or 'wrong' or whether the puppet 'said it right'. So actually, the name *Truth* Value Judgment Task is slightly misleading there. It is more of a *Truth-and-Felicity* Judgment Task.

was that as these notions also include felicity, more attention would be paid by the participants to the fit of the target sentence to the question. In order to isolate the effect of the wording of the task, the control experiment (Experiment 11) was an exact copy of Experiment 10, with the only difference that participants were instructed to judge whether Speaker B's answer was right or wrong, instead of true or false.

4. Experiment 11: Controlling for the wording of the task: true/false vs. right/wrong

4.1 Setup, items, design, participants and procedure

The setup, items, design and procedure of Experiment 11 were identical to Experiment 10, except for the fact that participants were instructed to judge whether Speaker B's answer was right or wrong, and the buttons 'true' and 'false' were replaced by 'right' and 'wrong'. The instruction that participants should ignore strangeness or ugliness of the answer was removed from the instructions.

43 participants filled out a web-based questionnaire (WWSTIM, Veenker 2000). None of them had participated in Experiment 10 or had any prior knowledge of the topic. All of them were adult native speakers of Dutch and most of them were students or had a university degree. Most participants completed the experiment within 15 minutes.

4.2 Results

The same exclusion criteria as in Experiment 10 were used for participants who were suspected of answering 'wrong' to a *How many* question because they considered an exact number to be required. 8 subjects were excluded based on these criteria. Another two participants were excluded because they

[4] An example of an experiment on the SI of *most* where the wording of the task made a big difference is a study by Papafragou & Schwarz (2006). They had a character in a story say *I'm going to color most of the star blue* and presented participants with an 'after' picture in which the whole star was colored blue. Then they asked participants *Did he do okay?* In 43% of the cases, the answer was 'no', indicating the SI. However in another (similar) experiment they asked *Did he do what he said?* and rejections dropped to only 10%.

scored less than 75% correct on the fillers. The results of the remaining 33 participants were included in the analysis.

The SI-rates of Experiment 11 are given in Table 3:

Table 3: SI-rates Experiment 11

Condition 1 (wh focus)	Condition 2 (wh non-focus)	Condition 3 (yes/no sent.)	Condition 4 (yes/no polar)
64%	48%	44%	38%

Again, a significant difference was observed between the two wh-conditions: Wilcoxon signed rank test (over participants): $z = -2.33$, $p = 0.010$ (one-tailed), so the effect of focus-sensitivity was replicated with this alternative task. The distribution of participants in the two wh-conditions was also similar to that of Experiment 10: 24% of the participants distinguished between the two conditions in the predicted direction, and almost none the other way around. See Table 4:

Table 4. Distribution of participants Experiment 11, wh-conditions

more SIs in foc	SIs in both cond	SIs in neither	more SIs in n-foc
24% (8)	39% (13)	33% (11)	3% (1)

The yes/no sentential condition (C3) again patterned with the non-focus wh-condition (C2), and not with the focus wh-condition (C1) (Wilcoxon signed rank test of C1 and C3: $z = -2{,}67$, $p = 0.004$ (one-tailed)). The yes/no polar condition (C4) again patterned with the yes/no sentential condition (C3-C4 did not differ significantly: $z = -1.41$, $p = 0.157$ (two-tailed)).

Overall, the results were similar to the results of Experiment 10. Between-subjects analysis did not reveal a significant difference between the two experiments in any condition (Mann-Whitney tests returned p-values >0.05 for all conditions).

4.3 Discussion

The focus-sensitivity of SIs was replicated in Experiment 11 with a different version of the experimental paradigm. Although the difference between the two focus conditions was slightly bigger in Experiment 11 than in Experiment 10 (16% vs. 10%), there was no significant difference between the two experiments. Therefore, it seems asking for true/false judgments or for right/wrong judgments has no effect on the size of the focus effect on SIs. The hypothesis that in Experiments 1-3 the difference between the focus

and non-focus conditions is smaller than predicted by the theory because participants ignore the fit to the question, therefore becomes less likely.

5. General Discussion Experiments 10 and 11

In this section I discuss the results of Experiments 10 and 11 in three parts. Section 5.1 focuses on the replication of the focus-sensitivity of SIs in the two wh-conditions, and the comparison to the TVJT-experiments with *or*. In 5.2 the results of the yes/no sentential condition are discussed, and in 5.3 I discuss the yes/no polar condition. While the latter condition was tested independently of the predictions of the focus-sensitive view of SIs, I argue this view might be extended to account for it.

5.1 Discussion of the wh-conditions

The results of conditions 1 and 2 of Experiments 10 and 11 showed that the focus-sensitivity of SIs extends to another scalar term in another structural position. As in the experiments with *or*, the difference was not black-and-white as predicted by the theory, but gradual. In section 5 of chapter 3 I explored three possible explanations for why this was the case in the *or*-experiments. The first explanation was that participants might have been guessing in the non-focus condition. This explanation also seems to be viable here, as the percentages in the non-focus wh-conditions in both experiments were close to 50% (resp. 42% and 48%). However, in the current experiments the percentages in the focus condition were also reasonably close to 50% (52% and 64%). Therefore, I will consider the hypothesis that participants were guessing in both conditions in 5.1.1 below. There I will also discuss the problem of a carry-over effect between conditions, of which the results suggest it was present.

The second explanation for the small difference, proposed in section 5.2 in chapter 3, was that there was a mismatch between the implicit question that was triggered by the description of the actual situation and the explicit question that was asked by the speaker. I will discuss whether this possibility extends to Experiments 10 and 11 in section 5.1.2. I argue that the setup of the stories increased this effect in the current experiments. There I also discuss a related problem, the fact that the experimental items involved a task that distracted away from the linguistic material.

The third explanation proposed in chapter 3 (in section 5.3), was that the focus manipulation could not completely overrule the sentence-final default position of focus. This explanation is rejected by the experiments in this

chapter, as we observed a comparable SI-rate in the non-focus condition to Experiments 1-3, even though the scalar term was no longer in the default focus position at the end of the sentence, but sentence-initial.

5.1.1 Chance-performance and carry-over

In section 5.1 of chapter 3 I already pointed out that we have to look at the distribution of participants over the possible types of behavior to determine if the SI-rates are due to guessing. This distribution is given in Table 5 for Experiment 10.

Table 5: distribution of participants over behavior Exp. 10

behavior	expected (guessing)	observed (wh foc)	observed (wh non-f)
2/2 'false'	5	15	9
1/2 'false'	10	1	7
0/2 'false'	5	14	14

For both the wh focus and the wh non-focus condition, the distribution was significantly different from the expected distribution based on guessing behavior (focus: $\chi^2 = 26.20$, $p < 0.001$, non-focus: $\chi^2 = 10.20$, $p = 0.006$). So as in Experiments 1 and 2, the individual behavior of participants in Experiment 10 did not reflect guessing.

However, the observed bi-modal distribution in the wh focus condition (with only one participant judging 1/2 items false), could be reflecting a type of guessing behavior I discussed in section 5.1 of chapter 3: making an initial guess and then sticking to it. This possibility is also reflected by the distribution of both wh-conditions of Experiment 11, given in Table 6.

Table 6: distribution of participants over behavior Exp. 11

behavior	expected (guessing)	observed (wh foc)	observed (wh non-f)
2/2 'false'	8.25	21	13
1/2 'false'	16.5	0	6
0/2 'false'	8.25	12	14

Again, the distributions did not reflect individual guessing, but participants were very consistent, which could mean they were sticking to an initial guess. However, in that case the number of participants that consistently said 'true' should be comparable to the number that consistently said 'false'. Although this matches the distribution in the focus condition of Experiment

10 and the non-focus condition of Experiment 11, it does not seem to match with the distribution of the non-focus condition of Experiment 10 (which is not clearly bi-modal) and of the focus condition of Experiment 11 (in which almost twice as many participants said 'false' 2/2 than 0/2 times).[5] I conclude that the data in general do not match a guessing pattern.

However, in Experiments 10 and 11 participants were very likely to answer all target items of *all four* conditions the same. In Experiment 10, 18 out of 30 participants (60%) answered all 8 target items the same, and in Experiment 11 this was 18 out of 33 participants (55%), substantially decreasing the difference between the conditions compared to the experiments with *or*. This points to the possibility that there was a lot of carry-over between the conditions, which is not surprising if we look at the large number of conditions (4) and the small number of fillers (8). This could also explain why the difference was smaller in Experiments 10 and 11 than Experiments 1 and 2. Experiments 10 and 11 contained 8 test items (2 per condition) and 8 fillers, where Experiments 1 and 2 contained 6 test items (3 per condition) and 14 fillers.

5.1.2 Inconsistency implicit question and explicit question and non-linguistic task

In 5.2 of chapter 3 I argued that due to the description of the actual situation in the story, the double-focus *Who found what?* question might have been triggered, leading to a mismatch with the focus-determining question asked by the speaker in the items. I argue this problem might have been even bigger here as the story very explicitly provided a list. Reconsider (5):

[5] In section 5.1 of chapter 3 I addressed this possibility by looking at the first item data. However, due to the higher number of conditions in Experiments 10 and 11, the first item of a certain condition that participants were presented with was often preceded by several other test items (of other conditions). Considering only the very first test items participants were presented with, as I also did for Experiments 1 and 2, gives too few data points to be informative. For instance, the 30 participants and 4 conditions in Experiment 10 gives an average of 7.5 data points per condition if we only consider the first test items.

(5) Condition 1 (wh focus)
Five people were present at the get-together of the Celtic language studies program at the University. Several drinks were available.
Sander is a student. He drank beer.
Tom is also a student. He also drank beer.
Eric is a professor. He drank wine.
Martin is a professor too. He drank apple juice.
Frans is a student. He drank beer.

A: "How many of the students drank beer?"
B: "Most students drank beer."

Due to the fact that the story is a list of people and beverages they chose, the *Who drank what?* question becomes very salient. Relative to this question, both subject and object are information focus, nullifying the critical manipulation due to A's question. This could explain why such a small difference was observed.

Another explanation related to this is the possibility that the items in Experiments 10 and 11 involved a task that distracted away from the linguistic material in the items. Due to the use of *most*, the task in Experiments 10 and 11 involved a lot of counting. All the stories were lists of people and the objects with which they interacted. The introductory sentence also always contained numbers (like *five* in (5)). In both conditions, to check a target sentence like *Most As did B*, participants had to count the number of As that did B, and see whether it exceeded the number of As that did not do B. All the items (including the fillers) contained *most*, (or *the minority of*), so all of them required counting.[6] This task might have led to a strategy in which participants upon encountering another test sentence with *most* simply started counting or checking the proportion of As that did B, without paying much attention to the question, which was the critical manipulation. Crucially, the question was not part of the situation which had to be counted to determine whether the target sentence was true. This could explain why more than half of the participants did not distinguish between any of the conditions in Experiments 10 and 11: they paid no attention to the question and were simply checking the *most* statement with the situation.

[6] It could be argued that in the test items, no counting was needed as all of the As did B. Nevertheless, it was still required to check all the As and keep score of which proportion of them did B.

5.2. Discussion of the yes/no-sentential answer condition

As predicted, the yes/no sentential condition patterned with the wh non-focus condition. Above I discussed a couple of explanations why the difference between the two wh-conditions was not bigger. However, the yes/no sentential condition also still yielded a higher number of SIs than predicted by the theory. The first obvious explanation for this is carry-over between the conditions. However, in this section I will discuss another possibility.

Let us look more closely at the assumption that answers to yes/no-questions are as a whole focus-less, which I took as a working assumption. This assumption might have been a bit too hasty. Jackendoff (1972) claims that questions themselves also have a focus structure. For wh-questions, this structure is relatively simple: the wh-phrase is the focus, and the rest of the sentence is the background. This follows straightforwardly from the generalization that was already referred to in chapter 2, that question and answer share the same focus structure (Rooth, 1996). However, with yes/no-questions, things are a bit more complicated. According to Scharten (1997), there are two types of yes/no-questions: neutral and topicalized. Neutral yes/no-questions are the type of questions in which no part has any special intonation, and which cannot be rephrased as wh-questions. An example of a neutral yes/no-question is (15):[7]

(15) Did you clean up your room?

Topicalized yes/no-questions on the other hand, are actually wh-questions in disguise. A part of the question receives extra stress, and that part corresponds to the wh-phrase in the corresponding wh-question. Scharten gives example (16), and example (17) is from Jackendoff (1972):

(16) Did JOHN take your books?
≈ Who took your books?

(17) Did Maxwell kill the judge with a HAMMER?
≈ What did Maxwell kill the judge with?

These topicalized yes/no-questions clearly have a focus structure, the same as their wh-equivalents. In (16), *John* is the focus, and in (17) *a hammer*.

[7] Of course, stress can change this neutral yes/no-question into a topicalized yes/no-question. But with a normal stress pattern, this question is a good example of a neutral yes/no-question.

It therefore seems reasonable to assume that a sentential answer to a topicalized yes/no-question, like an answer to a wh-question, inherits the focus structure of the question, e.g. as in (18):

(18) A: "Did Maxwell kill the judge with a HAMMER?"
B: "(Yes,) he killed him with a hammer$_F$."

If this is the case, we can no longer be sure that the sentential answers condition 3 were as a whole focus-less. As the dialogues were presented in written form, participants could have superimposed all kinds of stress patterns on the yes/no-question, possibly turning it into a topicalized yes/no-question, as in (19)-(22):

(19) Did most students drink BEER?

(20) Did MOST students drink beer?

(21) Did most STUDENTS drink beer?

(22) Did most students DRINK beer?

This means we can no longer assume *most* was non-focus in the answer. If for instance participants read the question as in (20), this would have caused *most* in the answer to be the focus, and the SI is predicted to arise. As all but two target sentences in the experiment (of both test items and fillers) contained *most*, it is not even unlikely that participants read the question with extra emphasis on *most*, so as a topicalized yes/no-question. This could explain why the difference between the yes/no sentential answer condition (C3) with the focus wh-condition (C1) was reasonably small. I now turn to the yes/no-questions themselves, as the analysis of topicalized yes/no-questions opens the possibility that the focus-sensitivity of SIs extends to questions.

5.3 Discussion of the yes/no polar condition

The yes/no polar condition (C4) returned a comparable SI-rate to the yes/no sentential condition (C3), indicating that the SI was calculated in the yes/no-question to the same extent as in the declarative answer. The most obvious explanation is again that there was carry-over from the declarative conditions. In section 3.5 of chapter 3 I referred to another experiment in the paper of Noveck et al. (2002) referred to in 2.3 above. There, SIs in

questions went up from 31% to 81% when they were presented after a declarative condition. Similarly, the SI-rate in the yes/no polar condition could have been assimilated to the SI-rates of the (non-focus) declaratives in this experiment.

However, if we adopt the view proposed in the previous section that questions themselves also have a focus structure, the question arises whether SIs could arise in the focus part of the (topicalized yes/no-) question. The proposals of Van Rooij and Van Kuppevelt do not address this issue, as their accounts are based on how answers are interpreted. However, intuitively, it is not unthinkable that (20) above is interpreted as *Did most but not all students drink beer?* which would lead to a 'no' answer in C4. This might explain why we observed a reasonable number of SIs in this condition.

In Zondervan (2010) I suggested a simple experiment with *or*, in which stress is used to turn the yes/no-question into a topicalized one with a focus structure, see (23) and (24):

(23) Focus condition:
Story: John has a and b.
Question: Does John have A OR B? yes/no

(24) Non-focus condition:
Story: John has a and b.
Question: Does JOHN have a or b? yes/no

In the topicalized yes/no-question (23) *a or b* is the focus of the question. If the focus-sensitivity of SIs extends to questions, we would expect an SI here, yielding the reading *Does John have A or B but not both?* Therefore, after a story in which it is described that John has A and B, we would expect participants to answer this question with 'no'. In the topicalized yes/no-question in (24) however, *John* is the focus and no SI should arise for *a or b*, resulting in a 'yes' answer. However, a structural investigation of the focus-sensitivity of scalar terms in questions falls outside the scope of this dissertation, so I leave it as a suggestion for further research.

6. Conclusion

In this chapter the scope of the investigation of focus-sensitivity of SIs was extended. I considered whether the effect of focus on SIs could be replicated for another scalar term in another scalar position. Although the differences were small, there was a significant difference between the focus and non-focus wh-conditions. The small differences might be due to a number of

factors. An important factor was probably carry-over between the conditions due to the high number of conditions and the small number of fillers. Furthermore, I pointed out two factors that might have decreased the effect of the critical manipulation: the possibility that the stories made the double-focus question (e.g. *Who drank what?*) salient and the task of counting that might have moved participants' attention away from the linguistic material which contained the critical manipulation.

The results also confirmed the prediction of the focus-sensitive account for scalar terms in sentential answers to yes/no-questions. These patterned with the non-focus wh-condition, which was predicted on the view that they are focus-less. The possibility of topicalized yes/no-questions might explain why still a reasonable number of SIs was observed in this condition. The possibility that participants considered the yes/no-questions to be topicalized can also account for why the scalar terms in the yes/no-questions themselves triggered almost as many SIs as scalar terms in their sentential answer counterparts. Based on this I argued that it is possible that the focus-sensitivity of SIs extends to questions, a hypothesis that goes beyond the scope of this dissertation.

In the next chapter I return to the scalar term *or*. So far, I have adopted the view that the exclusive reading of *or* comes about by an SI. I took the Gricean view of how the SI of *or* is derived (and the implementation of the Gricean view in terms of exhaustivity by Van Rooij (2002)). Although the view that the exclusive reading of *or* comes about by an SI is the dominant view in the literature, it has often been noted the Gricean view is problematic due to a paradox in assumptions about speaker expertise. In the next chapter I discuss the problem and I present additional data from Experiment 4 that addressed this issue. I also present the theoretical solutions to this problem that have been proposed in the literature and consider whether they can account for the paradox.

CHAPTER 8: THE SPEAKER-EXPERTISE PARADOX FOR THE SI-VIEW OF *OR*

1. Introduction

In chapters 3-6 of this dissertation I presented experimental support for the claim that when a constituent of the form *A or B* is part of the information focus of the sentence, it is more likely to be interpreted as *A or B but not both* (exclusive-*or*) than when it is part of the background. So far, I have adopted the dominant view in the literature that this reading of *or* comes about by the SI that the stronger scalar alternative *A and B* does not hold. However, it has often been noted in the literature that the Gricean SI-view is problematic for *or*. The crucial issue is that the Gricean view requires the assumption that the speaker is well informed. However, the use of *or* (at least in episodic sentences) also typically triggers the inference that the speaker is not completely informed. This discrepancy is often considered to be problematic for Gricean SI-view of *or* (see e.g. Geurts 2006). In the first part of this chapter I consider this problem in more detail, and I present additional data from Experiment 4 (discussed in chapter 4) which address this problem.

In the second part of this chapter I introduce four recent theories of the SI of *or* which might account for this problem. The first two theories, by Sauerland (2004) and Van Rooij & Schulz (2004), adhere to a derivation of SIs by (some form of) Gricean reasoning, but provide a more sophisticated mechanism for including the assumption of speaker expertise. Both derive SIs in a two-step procedure. These views will be discussed in section 4. In section 5 I turn to two theories of SIs that abandon the view that SIs come about by Gricean reasoning and claim SIs are derived in the grammar. The first view is the account of Fox (2007) who proposes SIs are derived by an exhaustivity operator in the grammar which applies to the whole sentence. The second view is the approach of Chierchia (2004, 2006) and Chierchia, Fox and Spector (2008), who propose SIs are derived by an exhaustivity operator in the grammar which can be applied locally, during compositional semantics. As these views do not rely on Gricean reasoning to derive the SI, they might avoid the problem of speaker expertise. However, I argue all four accounts cannot solve the problem of speaker expertise and at the same time account for the focus-dependency observed in the previous chapters.

Three of the four accounts that will be discussed rely on some form of exhaustivity, which I will consider in the next chapter by presenting additional data on the relation between focus, exhaustivity and SIs from Experiments 5-7.

2. The Speaker Expertise Paradox of *or* for the 'simple Gricean view'

As discussed in chapter 1, SIs as in (1) below are traditionally considered to be a type of conversational implicature, which were introduced by Grice (1967). The Gricean view of conversational implicatures is that they come about by (some version of) a nonce derivation like (2). See section 2 of chapter 1 for a more elaborate description of the Gricean view.

(1) Laura found a crab or a starfish.
SI: It is not the case that Laura found a crab and a starfish.

(2) i. The speaker used the scalar term *or*.
ii. The speaker could have uttered the same sentence with the scalar term *and* instead of *or*, which would have been stronger / more informative (because the sentence with *and* entails the sentence with *or*).
iii. The sentence with the stronger scalar term *and* would have also been relevant.
iv. The speaker is trying to be as informative as possible (she is obeying the Maxim of Quantity).
v. Apparently, the speaker does not have evidence for the sentence with *and*.
vi. The speaker is well informed.
vii. Therefore, it is likely that the speaker considers the sentence with *and* to be untrue.

When discussing the defaultist/contextualist debate in chapter 5, I already pointed out that some authors, although adhering to the Gricean program, have proposed SIs do not necessarily rely on a nonce derivation like (2), as they might be generated by default (e.g. Levinson, 2000). However, the processing findings in the literature as well as in this work have rendered default generation unlikely (see section 6 of chapter 6). Therefore, I will focus on the view that SIs come about by (some version of) the nonce derivation. I will call this the *simple Gricean view*, to be contrasted with more complex Gricean mechanisms of deriving the SI which will be discussed in section 4 below. I will focus on step (vi) of the nonce derivation: the assumption of speaker expertise.

Step (vi) in the nonce derivation (2) above is the assumption that the speaker is well informed. Specifically, the addressee has to assume the speaker is at least well informed about the stronger scalar alternative. Otherwise, the addressee cannot conclude that the speaker probably considers that alternative to be untrue, and the SI would not be derived.

However, a number of authors have noticed that this assumption leads to problems with the scalar term *or* (e.g. Geurts 2006). Next to the *not and*-SI, *or* also typically gives rise to what Gazdar (1979) calls *Clausal Implicatures*, also frequently called *Ignorance Implicatures*. These are the inferences that when a speaker utters something of the form *A or B*, the addressee concludes the speaker does not know the truth value of *A* and does not know the truth value of *B*. Although Gazdar proposed a separate rule to derive these Clausal Implicatures, in principle they can be derived by a nonce derivation too, given in (3):

(3) i. The speaker used the expression *A or B*.
ii. The speaker could have uttered the same sentence with *A* instead of *A or B*, which would have been stronger / more informative (because the sentence with *A* entails the sentence with *A or B*).
iii. The sentence with the stronger *A* would have also been relevant.
iv. The speaker is trying to be as informative as possible (obeying the Maxim of Quantity).
v. Apparently, the speaker does not have evidence for the sentence with *A*.

However, we cannot take step (vi) and (vii) as in (2), because if the speaker considers A to be false, and she considers *A or B* to be true, she has to consider B true. However, then the Maxim of Quantity would have ordered her to utter *B* instead of *A or B*. Moreover, we can apply the reasoning in (3) to *B* too, so steps (vi) and (vii) give the result that the speaker considers both *A* and *B* to be false, which leads to a contradiction with the assertion *A or B*. Hence, we cannot take step (vi) and (vii) and we are left with the inferences that the speaker does not know the truth value of *A* and does not know the truth value of *B*. However, to derive the SI of *A or B* (that it is not the case that *A and B*), we have to take step (vi) and (vii) (see (2)). Otherwise, we can only derive the weak SI that the speaker does not know whether *A and B* is true. So the simple Gricean view leads to counterintuitive assumptions of speaker expertise: the addressee has to assume that the speaker is well informed about the truth value of *A and B* (otherwise the strong *not and*-SI cannot be derived), but at the same time she has to assume the speaker does not know the truth value of *A*, nor the truth value of *B*. I will call this the *Speaker Expertise Paradox (SEP)* from now on, see (4):

(4) **Speaker Expertise Paradox (SEP) of *or* for the simple Gricean view**
In order to derive the Clausal Implicatures and the Scalar Implicature of *A or B* on the simple Gricean view, it has to be the case that the addressee assumes that the speaker is in an epistemic state in which she knows the truth value of *A and B*, but not the truth value of *A*, nor the truth value of *B*.

The epistemic state described in (4) is possible. Consider for instance a situation in which you are standing in front of a sandwich place and your friend John tells you he is going to get a sandwich and he is not sure which one of his two favorite sandwiches (say, the tuna sandwich and the cheese sandwich) he is going to get. A minute later you see him walking out of the sandwich place carrying one sandwich wrapped in paper. He waves at you and takes off. At that point, you might be in the epistemic state in which you know that it is probably true that he bought a tuna sandwich or a cheese sandwich, you do not know which one he bought, but you do know that he probably did not buy both.

Even though a situation like the one just sketched is possible, the sheer number of words I needed to set up the situation indicates that this is a very special situation, and that it is highly *unlikely* that someone is in such an epistemic state. Therefore, it is also very unlikely that an addressee will make the assumption that a speaker is in such an epistemic state. However, in natural language, *or* is interpreted as exclusive-*or* very frequently (see the results in the previous chapters of this work), much more frequently than a situation like the one sketched above is likely to occur.[1] Notice that even if we do not consider Clausal Implicatures to be derived by a nonce derivation like (3), but assume they come about in some other way, the problem for the simple Gricean view of SIs stays. To derive the SI by Gricean reasoning, we have to make an assumption about speaker expertise (that the speaker knows *A and B* is false) which is unlikely to co-occur with the Clausal Implicatures (that the speaker does not know the truth value of *A* nor the truth value of *B*). So even if we take the Clausal Implicatures to come about in some other way, the SEP stays.

[1] Of course, the assumption that the speaker knows that *A and B* is false can also be the result of world knowledge. For instance, the addressee only has to assume that the speaker has the common sense belief that a person cannot be at two places at the same time, to know that when a speaker utters *John is at home or in the office*, he knows he is not at home *and* in the office. (This point is also brought up by Geurts (2006)). Leaving aside the issue whether an SI will arise here anyway, even if we disregard these situations, I believe exclusive-*or* is much more frequent than the specific assumption of speaker expertise that is needed on the simple Gricean view.

In the next section I will (re)consider Experiment 4, which tested whether participants indeed calculate Clausal / Ignorance Implicatures and SIs at the same time. The rationale behind this is that if a participant calculated the SI of *A or B* in a certain environment, on the simple Gricean view she must have considered the speaker to be informed about the truth value of *A and B*. If however the participant at the same time made Clausal / Ignorance Implicatures, on the simple Gricean view we have to assume she considered the speaker to be in the highly unlikely epistemic state described by the SEP. That would render it very unlikely that the simple Gricean view is the right account of the SI of *or*.

3. Experiment 4: Testing speaker expertise

3.1 Setup and items

Experiment 4, already partly presented in chapter 4, was set up in the *Possible World Judgment Task (PWJT)* paradigm, in which participants were presented with a story containing a dialogue, such as (5):

(5) Example item Experiment 4
Marieke told her mother that Laura went searching for marine animals on the beach yesterday, and that her father had told her that if she would find at least two marine animals, she would get to stay up late that night.
The mother said: "Oh, and what did Laura find?"
Marieke answered: "She found a crab or a mussel."

SIs were assessed by asking participants to answer questions like (6):

(6) Do you think it is possible that Laura found both a crab and a mussel? yes / no

A 'no' answer to this question indicated that the participant interpreted *or* as exclusive-*or* (*a crab or a mussel but not both*), which is traditionally considered to be the result of an SI. In chapter 4 I considered how the contextual manipulations of information focus and alternative relevance in Experiments 4 affected the answers to questions like (6).

Every item in Experiment 4 also contained a question which measured speaker expertise, for instance (7) for (5) above:

(7) Do you think Marieke knows exactly what Laura found? yes / no

This question does not directly ask for the Clausal / Ignorance Implicatures (*IgnIs* from now on). The questions asking for IgnIs would have been something like: *Do you think Marieke knows whether Laura found a crab?* and *Do you think Marieke knows whether Laura found a mussel?* but I felt those questions would have been too confusing for the participants. It would make participants wonder why one of the two disjuncts was chosen and not the other, possibly triggering additional assumptions about the situation (e.g. asking about the crab might be taken to indicate that finding a crab was more important than finding a mussel, or that in fact Laura found a crab and not a mussel). A 'no' answer to (7) indicated that the speaker was not considered to be completely informed, and the most likely thing to have caused this assumption is the speaker's use of *or*, which is normally associated with the IgnIs. To control for other factors which might have caused participants to consider the speaker not to be fully informed, I included a control condition without *or*, which I present below (condition 4).

However, even if questions like (7) were not a good measure of IgnIs, the simple Gricean view is still in trouble when SIs are observed while they were answered negatively. Such a result indicates the SI is calculated despite the fact that the speaker was not fully informed. This raises (a version of) the Speaker Expertise Paradox, which is problematic for the simple Gricean view.

Every item contained both an SI-question like (6) above and a speaker expertise question like (7). This way we could investigate whether participants made both inferences at the same time. Additionally, a comprehension question was asked to make sure participants read the critical parts of the stories. Each story was therefore followed by three questions, exemplified in (8). The material between parentheses was not present in the items.

(8) Question 1 (comprehension-question):
What would Laura be allowed to do if she would find a crab or a mussel?...

Question 2 (speaker-expertise question):
Do you think Marieke knows exactly what Laura found? yes / no

Question 3 (SI-question):
Do you think it is possible that Laura found both a crab and a mussel? yes / no

I chose this order of questions as I felt it was the most natural. Furthermore, as the SI-question followed the IgnI-question, it allowed us to assess whether

SIs were calculated in absence of the assumption of complete speaker expertise.

As discussed in chapter 4, the experiment also tested the effect of focus and alternative relevance on SIs. Therefore, there were three test conditions in which focus and alternative relevance were manipulated, repeated from chapter 4 in (9)-(11). The critical manipulations are highlighted in boldface.

(9) Condition 1: [-alternative relevance, +focus]
Marieke told her mother that Laura went searching for marine animals on the beach yesterday, and that her father had told her that **if she would find a crab or a mussel**, she would get to stay up late that night.
The mother said: "Oh, and **what did Laura find?**"
Marieke answered: "She found a crab or a mussel."

(10) Condition 2: [+alternative relevance, +focus]
Marieke told her mother that Laura went searching for marine animals on the beach yesterday, and that her father had told her that **if she would find at least two marine animals**, she would get to stay up late that night.
The mother said: "Oh, and **what did Laura find?**"
Marieke answered: "She found a crab or a mussel."

(11) Condition 3: [-alternative relevance, -focus]
Marieke told her mother that Laura **and Barbara** went searching for marine animals on the beach yesterday, and that their father had told them that **the one who would find a crab or a mussel**, would get to stay up late that night.
The mother said: "Oh, and **who of them found a crab or a mussel?**"
Marieke answered: "Laura found a crab or a mussel."

The conditional in the first sentence was varied to manipulate relevance of the stronger scalar alternative (*Laura found a crab and a mussel*), and the question in the dialogue manipulated the focus structure of the target sentence. The results of these manipulations were discussed in chapter 4, and are not relevant for our current purposes.

There were no predictions for an effect of the manipulations on the speaker expertise question. If we consider IgnIs to be derived by a nonce derivation, we expect them to be dependent on alternative relevance. However, the stronger alternatives in the nonce derivation by which the IgnIs are derived (see (3) above) are the disjuncts (e.g. *Laura found a crab*

and *Laura found a mussel*). For the conditionals in (9)-(11) it does not matter whether these were used in the target sentence instead of the disjunction, so all three conditions were [-alternative relevance] for the IgnIs. There were also no predictions about the effect of focus on the IgnIs. The only question we are interested in here is whether it occurred that SIs were calculated while the IgnIs were also made, so whether questions 2 and 3 of (8) were both answered with 'no', irrespective of condition.

To control for other reasons than the use of *or* which might have caused participants to consider the speaker not fully informed (and to control for other reasons which might have caused the participants to consider the *A and B* situation impossible or implausible), I introduced a control condition. This condition was identical to condition 3 but contained an answer in which the VP containing *A or B* was elided, see (12):

(12) Condition 4: (control)
Marieke told her mother that Laura and Barbara went searching for marine animals on the beach yesterday, and that their father had told them that the one who would find a crab or a mussel, would get to stay up late that night.
The mother said: "Oh, and who of them found a crab or a mussel?"
Marieke answered: "Laura did."

If participants considered the speaker not to be completely informed in the test conditions for some other reason than the use of *or*, the same story in condition 4 should also return many 'no' answers on the speaker expertise question. Therefore, by comparing the speaker expertise answers of C3 and C4, we get a clear view of what the contribution of the IgnIs of overt *or* is to the overall ignorance score. However, as I pointed out above, the simple Gricean view is already in trouble when SIs are observed while the speaker expertise question is answered negatively, irrespective of what caused the assumption of ignorance.

3.2 Design, participants and procedure

Four lists were created, each with 3 items per condition. 50 participants participated in the experiment, which was a web-based questionnaire (WWSTIM, Veenker 2000). See sections 4.3 and 4.4 of chapter 4 for a more elaborate description of the design, participants and procedure.

3.3 Results

The results of Experiment 4 are given in Table 1, where the percentages of 'no'-answers are given, which indicate IngIs and SIs:

Table 1: IgnI- and SI-rates Experiment 4

	Condition 1 [-alt +foc]	Condition 2 [+alt +foc]	Condition 3 [-alt -foc]	Condition 4 [control]
IgnI	93%	96%	92%	47%
SI	67%	72%	63%	27%

Table 1 shows that all three test conditions returned high IgnI-rates and reasonably high SI-rates. This indicates that in many cases both the IgnI and the SI were drawn at the same time. This is confirmed when we look at the distribution of answer combinations on the speaker expertise questions and SI-questions, given in Table 2.

Table 2: Answer combinations given in Experiment 4

	Cond. 1 [-alt +foc]	Cond. 2 [+alt +foc]	Cond. 3 [-alt -foc]	Cond. 4 [control]
Q2: 'no', Q3: 'no' (+IgnI, +SI)	61%	69%	58%	15%
Q2: 'no', Q3: 'yes' (+IgnI, -SI)	32%	27%	34%	32%
Q2: 'yes', Q3: 'no' (-IgnI, +SI)	6%	3%	5%	11%
Q2: 'yes', Q3: 'yes' (-IgnI, -SI)	1%	1%	3%	41%

The first row represents in how many cases in a condition for one item both the speaker expertise question (Q2), and the SI-question (Q3) were answered with 'no', indicating that both inferences were drawn. We see that in all three conditions in which the target sentence contained *or* (C1-C3), in more than half of the cases (resp. 61%, 69% and 58%) the SI was calculated while the speaker was not considered to be totally informed. This is problematic for the simple Gricean view of SIs, as the Speaker Expertise Paradox arises.

When we go back to Table 1, we see a clear pattern for the speaker expertise data. When *or* is absent (C4), participants seemed to be guessing about whether the speaker was informed or not (hence the score close to 50%). This is not surprising, as there was no indication in the story whether the speaker was informed or not, and the speaker did not specify what Laura

found. However, when *or* is present, participants almost always judged the speaker not to be totally informed, irrespective of focus or alternative relevance (the differences between the three test conditions (C1-C3) were not significant). The difference between C3 and C4 was significant: Wilcoxon signed rank test: $z = -5.204$, $p<0.001$ (two-tailed). Therefore, the data show that *or* was indeed responsible for the high rate of 'no' answers to the speaker expertise question. Although no effect was predicted and it is not crucial for the observation that SIs and IgnIs were calculated at the same time, it is interesting to see that IgnIs seem to be insensitive to focus. I will return to this in the discussion.

3.4 Conclusions and discussion

The data show the problematic situation for the simple Gricean view was observed. In more than half of the cases participants considered the speaker not to be totally informed, yet *or* received an exclusive interpretation. However, the attentive reader will remember that in section 7 of chapter 4 I raised two possible problems for the experimental paradigm that was used, the PWJT. I argued that at least one of them was a very plausible explanation why the differences between the SI-rates of conditions 1, 2 and 3 was so small. However, here the differences between the conditions are not relevant for our goal of determining whether SIs and IgnIs co-occur. But let us nevertheless consider the alternative explanations. After all, if the SI-data were affected by the paradigm in some way, we have to check whether this does not interfere with our conclusions here. I will argue that it does not.

The first explanation, presented in section 7.1 of chapter 4, was that a Manner Implicature interfered. This explanation is based on the possibility of a shorter answer in condition 3 (the term answer given in condition 4). As this shorter answer was possible, participants might have considered the repetition of the VP containing *or* to be somehow marked. Participants might have taken from this marked way of answering the question that the speaker considered the object to be relevant, or wanted to indicate she did not know exactly what was found. This might have increased 'no' answers to the speaker expertise question in condition 3. However, it seems unlikely that a Manner Implicature was responsible for both an increase in SI-rates and an increase in IgnI-rates at the same time, exactly because of the Speaker Expertise Paradox. If the hearer assumed the speaker wanted to communicate something extra by repeating the disjunction, it is unlikely that the hearer will assume the speaker wanted to communicate that she is ignorant about the truth values of the disjuncts (the IgnI), but at the same time she wanted to communicate that she does know *A and B* is false (the

SI). Anyway, irrespective of whether the SI-rates or the IgnI-rates in condition 3 were overestimated because of a Manner Implicature, the data of condition 1 and condition 2 cannot be explained by this, so they remain a problem for the simple Gricean view.

The second explanation I proposed for the small differences between the SI-rates of conditions 1-3 was that the experimental question overruled the critical manipulations (see 7.2 of chapter 4). As the experimental question explicitly asked about the *A and B* situation, this provided alternative relevance in all conditions. Also, as it was always about the direct object, it overruled the focus manipulation, turning all conditions into focus conditions. Finally, it was the most recent issue when participants interpreted the target sentence, instead of the critical manipulations. I argued this is a very plausible explanation for the fact that the differences between the SI-rates were small. In the same way, one could argue that by explicitly questioning the informedness of the speaker, we made it salient that she used *or* and therefore might not be totally informed. This could indeed have boosted the IgnI-rates and made them similar in all three conditions. However, as I pointed out above, it is not that important how the assumption of ignorance came about. What is important is that in more than half of the cases in all experiments, after participants answered 'no' to the speaker expertise question, they still interpreted *or* exclusively, judged from their 'no' answer at the consecutive SI-question. This is problematic for the simple Gricean view.

I conclude the criticism on how the simple Gricean view derives the SI of *or* is justified. Therefore, in the next section I discuss two proposals which account for the discrepancy between IgnIs and SIs, by providing a more sophisticated model of how speaker expertise enters the Gricean derivation. After that I also discuss two accounts that abandon the Gricean view altogether, and therefore might avoid the speaker expertise paradox. These theories claim the SI is derived through an exhaustivity operator in the grammar. As some aspects of the four theories that will be presented are quite technical, for each of them I discuss the mechanics first and I evaluate how they avoid the SEP in separate subsections.

4. The two-step Gricean accounts: Sauerland (2004) and Van Rooij & Schulz (2004)

In this section I will discuss two proposals that provide an alternative to the simple Gricean view presented above. These more complex Gricean views both propose a more sophisticated mechanism of adding speaker expertise.

Both rely on a two-step procedure to derive Clausal / Ignorance Implicatures and SIs through Gricean reasoning.

4.1 Sauerland (2004)

4.1.1 The mechanism

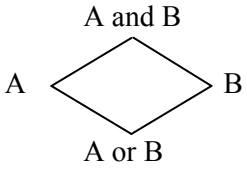
In order to account for the Clausal Implicatures and the SI of or, Sauerland (2004) proposed a system in which there are two types of implicatures, so called *Primary Implicatures*, which have weak epistemic force (to be explained below), and *Secondary Implicatures*, with strong epistemic force (that the speaker knows the stronger alternative is false). So far I have only considered the strong version of SIs, and described the weak Clausal Implicatures as the inference that the speaker *does not know the truth value* of the disjuncts. However, Sauerland proposes both SIs and Clausal Implicatures are first derived as weak implicatures, and he adopts a slightly different version of weak implicatures due to Soames (1982). On the description of weak implicatures used so far, the weak version of the SI of *A or B* would then be that the speaker does not know the truth value of *A and B*. That means she takes both *A and B* and ¬(*A and B*) to be possible. However, the version of Soames (1982) is that the speaker does not *have the positive belief* that *A and B* is true. This roundabout way of describing the speaker's beliefs is to avoid the inference from *does not believe* to *disbelieve*. Crucially, the speaker does not disbelieve the stronger alternative (as that is the strong implicature), but she does not have the positive belief that the stronger alternative is true. That could also mean she has no opinion about the stronger alternative (she neither believes nor disbelieves it). Hence, she takes it to be possible that ¬(*A and B)*. This version of the implicature however says nothing about the speaker's assumptions about the possibility of *A and B*.[2] This will turn out to be important later on.

The essential part of Sauerland's proposal is that the set of implicatures of a sentence is derived in two steps. First, the Primary (weak) Implicatures are derived, and then they can be turned into Secondary (strong) Implicatures. This can be done if an additional assumption is made, the assumption that the speaker knows the truth value of the stronger alternative. This is what Sauerland calls the *Epistemic Step*. However, there is a limitation to this step, which is that the Secondary Implicatures have to be consistent with the conjunction of the assertion and the set of Primary

[2] To see this, one has to realize that *it is possible that p* does not entail *it is possible that not p*, but that this is actually an SI (see Levinson 1983, p.140).

Implicatures. As Sauerland assumes Clausal Implicatures and SIs come about by the same mechanism, it follows that the disjuncts themselves also have to be scalar alternatives, so the set of alternatives for *A or B* looks like (13):

(13)



Notice that the set of alternatives is only partially ordered by entailment, as there is no entailment relation between *A* and *B*, hence the diamond shape of the scale. The primary implicatures of *A or B* are the ones given in (14):

(14) Primary Implicatures of *A or B*:
  a. The speaker does not have the positive belief that *A*.
  b. The speaker does not have the positive belief that *B*.
  c. The speaker does not have the positive belief that *A and B*.

The implicature in (14c) follows from (14a) and (14b). Notice that on the other paraphrase of weak implicatures (*The speaker does not know the truth value of* ...) this is not the case.

The Clausal Implicatures follow from taking together the assertion and these Primary Implicatures. If the speaker considers *A or B* to be true, and she considers it possible that *A* is false (14a), it follows that she has to consider *B* possible. Similarly, from the assertion and (14b) follows that the speaker considers *A* possible. Now, the question is which Primary Implicatures can be turned into Secondary Implicatures. The candidates are given in (15):

(15) Candidate Secondary Implicatures of *A or B*:
  a. The speaker knows that *A* is false.
  b. The speaker knows that *B* is false.
  c. The speaker knows that *A and B* is false.

The potential strong implicature (15a) is blocked by the assertion together with (14b), as from that followed that the speaker considered A possible. Similarly, (15b) is blocked by the assertion together with (14a). The strong implicature (15c) is consistent with the assertion and the Primary

Implicatures, so it can stay. This is how we derive the strong SI on Sauerland's system.[3]

Summing up, on Sauerland's analysis the weak Clausal Implicatures and the strong SI are the result of a two-step procedure, in which priority is given to weak implicatures. As a result, the epistemic step is only allowed to be taken for strong implicatures that do not contradict the assertion together with the set of all weak implicatures. Therefore, the epistemic step cannot be taken for the disjuncts, as excluding the one would be incompatible with the weak implicature of the other. However, it can be taken for the *and*-alternative, as there no contradiction arises.

4.1.2 Sauerland's model and the Speaker Expertise Paradox

Sauerland's system is an elegant solution for how the derivation of the strong SI does not lead to exclusion of the disjuncts. However, I argue that for SI-calculation it still runs into the Speaker Expertise Paradox. The question is why an addressee would ever take the Epistemic Step after the Clausal Implicatures have been derived by the first step of Sauerland's system. After the first step the addressee concludes the speaker has no positive knowledge that A is true nor that B is true and from this follows she has no positive knowledge that *A and B* is true. Knowing this, why would the addressee assume the speaker knows *A and B* is false? This assumption is still paradoxical. The only way around this problem is to assume that the Epistemic Step is taken by default. Sauerland's view is set in the neo-Gricean framework that (weak) SIs are generated irrespective of context, so a default epistemic step would reduce his framework to a default view of (strong) SI generation. We have seen this view is incompatible with the processing findings in the literature and this work (see chapter 6). Nevertheless, we could propose a contextualist version of Sauerland's proposal including a default epistemic step and make the ad-hoc claim the mechanism is only activated if the scalar term is in the focus. However, a similar two-step proposal was made by Van Rooij & Schulz (2004), building on the work of Van Rooij (2002) (discussed in chapter 2 of this thesis) and Schulz (2003). The advantage of this account over Sauerland's is that it is based on exhaustivity, which predicts the focus-sensitivity of SIs. I turn to this view now.

---

[3] Notice that this would not have worked with the other version of weak implicatures. If instead of (14c), the Primary Implicature would have been *The speaker does not know the truth value of A and B*, it would have followed that the speaker considers *A and B* to be possible, and (15c) would have been blocked.

4.2 Van Rooij & Schulz (2004)

4.2.1 The mechanism

Van Rooij & Schulz (2004) (VR&S) also derive the implicatures in a two-step procedure, in which the first step gives the Clausal Implicatures and the second step the SI. However, instead of the nonce derivation they introduce two exhaustivity operators (which they later combine into one). These operate on the semantics meaning of the sentence and give rise to the implicatures. Nevertheless, their account is still Gricean in nature as it formalizes Gricean reasoning about the speaker's knowledge. The first exhaustivity operator models Grice's maxims of Quality and Quantity-1 (that the speaker is being as informative as she has evidence for, corresponding to step (iv) of the nonce derivation), and the second operator adds speaker expertise (the crucial step (vi) of the nonce derivation). Similar to Sauerland's system, the second step (adding speaker expertise) takes the results of the first step into account. On the account of VR&S, the second step applies to the output of the first step, so competence is only added insofar as it is compatible with the speaker obeying Quality and Quantity. As their system is rather complicated, I present step 1 and step 2 in two separate subsections.

4.2.1.1 Step 1: eps_1

The first operator, called *eps*$_1$, accounts for Clausal Implicatures and was already proposed by Schulz (2003). It makes use of an information ordering between epistemic states and is defined as in (16), where P is the question predicate and A is the answer, $\Box$A means 'the speaker knows that A', s is an epistemic state, S is the class of states where the speaker is fully introspective and has true beliefs, and $\leq_\Box$ is the ordering.

(16) $eps_1^S(A,P) = \{s \in S | s \models \Box A \wedge [\forall s' \in S : s' \models \Box A \rightarrow s \leq_\Box s']\}$

The ordering $\leq_\Box$ is defined as in (17), where M stands for a model and R is a relation that represents the knowledge state of the speaker.

(17) $s_1 = \langle M_1, w_1 \rangle \in S$:
$s_1 \leq_\Box s_2$ iff $\forall v_2 \in R_2[w_2]\ \exists v_1 \in R_1[w_1]$
$V_1(P)(v_1) \subseteq V_2(P)(v_2)$

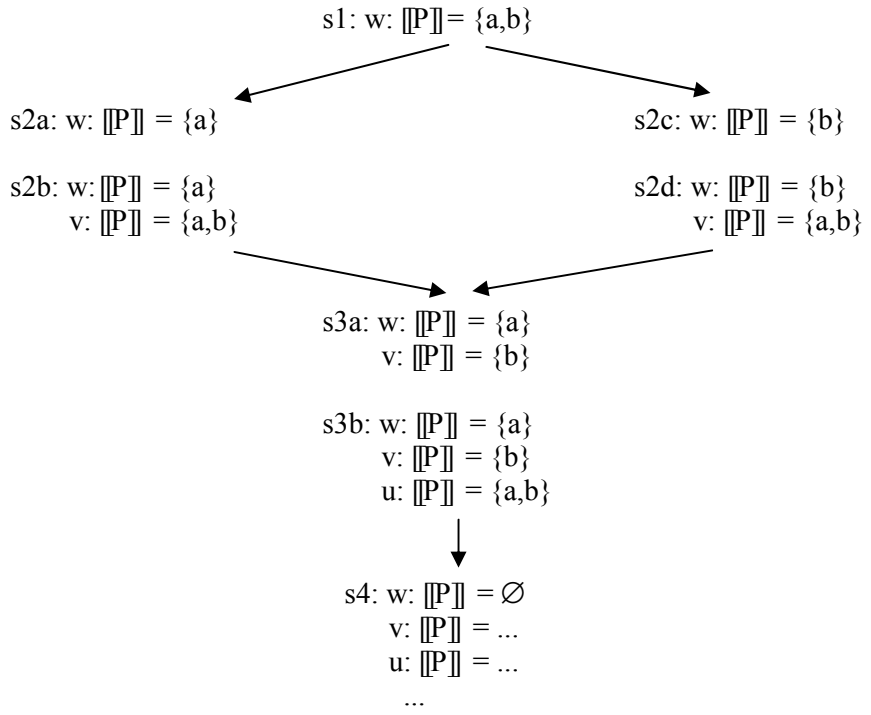
$s1 \cong_\Box s2$ iff $s_1 \leq_\Box s_2$ and $s_2 \leq_\Box s_1$

Eps_1 picks out the epistemic state of the speaker that is most minimal on $\leq_{\square}$, which is the epistemic state in which the speaker knows *least* about the extension of the question predicate P. The idea is that Quantity-1 commands the speaker to be as informative as possible, so eps_1 picks out the epistemic state in which the utterance was the most informative one that the speaker could have made. As a result, we can exclude epistemic possibilities in which the speaker knows more, because then she would have said so. Only positive information about P counts: knowing that an element is in the extension of P counts as knowledge, but knowing that an element is not in the extension of P does not count. Hence, the speaker knows less if she knows of fewer elements that they are in the extension of P. For instance, knowing that *a* is in the extension of P is more minimal than knowing that *a* and *b* are in the extension of P. This is expressed in (17) above, by saying that an epistemic state *s* is more minimal than *s'* iff for every epistemic possibility in *s'*, *s* contains an epistemic possibility where the extension of P is a subset of the one in *s'*. For the example just given, if the only epistemic possibility in *s'* is that *a* and *b* are in the extension of P, *s* is more minimal than *s'* if *s* contains an epistemic possibility that only *a* (or only *b* for that matter) is in the extension of P, as {a} is a subset of {a,b}.

VR&S illustrate the workings of eps_1 by providing a figure for the order that $\leq_{\square}$ gives to all possible epistemic states if the domain consists of only two elements, *a* and *b*. As I think their notation might be confusing, I provide a slightly different figure:[4]

[4] VR&S use set notation for the epistemic possibilities of the extension of P, while the extension of P itself would normally also be given in set notation. To avoid confusion over these different types of sets, I provide a figure in which the epistemic possibilities are presented as a list instead of a set, and I have numbered the different epistemic states.

Figure 1: ordering of eps_1 in a domain with two objects



This figure shows the ordering relation $\leq_{\square}$ on the different possible epistemic states (s1... s4), where s4 is the most minimal. The states are sets of epistemic possibilities, given in a list. If states are on the same level of the ordering, they are labeled *a*, *b* etc. (e.g. s3a $\cong_{\square}$ s3b). The ordering represents how little the speaker knows about the extension of P (where the speaker knows less if she knows of fewer elements that they are in the extension of P). This is reflected straightforwardly in the figure: In s1 the speaker knows that *a* and *b* are in the extension of P. In s2a-b the speaker knows that *a* is in the extension of P, and in s2c-d the speaker knows that *b* is in the extension of P. Therefore, s2a-d are more minimal than s1. Notice that s2a is not more minimal than s2b, (and s2c is not more minimal than s2d) as the negative knowledge in s2b (that *b* is not in the extension of P) does not count. In s3a-b, the speaker knows that at least one of *a* or *b* is on the extension of P, but not which one. Therefore, these states are more minimal than s2a-d, where the speaker is sure of one of the two. Finally, in s4, which represents all epistemic states which contain the epistemic possibility that the extension of

P is the empty set, the speaker does not know of *a* nor of *b* whether it is in the extension of P, and therefore these states are more minimal than s3a-b.

If we apply eps_1 to the simple semantic meaning P(b), we first select the states where the speaker knows that *b* is in the extension of P. Those states are s1, s2c and s2d. Of these states, s2c and s2d are the most minimal, so those are selected. Hence, the speaker considers it at least possible that *a* is not in the extension of P. Applied to a disjunctive meaning like P(a) ∨ P(b), we have to select from the states in which the speaker knows that *a* or *b* is in the extension of P, which are s1, s2a-d, s3a-b. The most minimal states of these are s3a and s3b, so eps_1 picks out those. In those states, the speaker does not know the truth value of the disjuncts, which captures the weak Clausal Implicatures. Furthermore, it follows that the speaker does not know whether P(a) ∧ P(b) is true, as this is different in s3a and s3b. This gives the epistemic weak SI that was proposed by e.g. Soames (1982): that the speaker does not have the positive belief that P(a) ∧ P(b).[5] So eps_1 can account for both types of weak implicatures (clausal and scalar) that arise from a disjunction. However, VR&S still have to derive the strong SI, which obviously does not follow from eps_1. Therefore, they introduce the second step.

4.2.1.2 Step 2: eps_2

To derive the strong SIs, the notion of *speaker competence* has to be added to the working of eps_1, similar to Sauerland's Epistemic Step. First, VR&S try to extend the notion of Groenendijk & Stokhof (1984) of *knowing whether* to achieve this. This notion entails that a speaker is competent if for every element that is in the extension of the question predicate P in the real world, she knows this is so. It could be proposed to apply eps_1 only to these states. However, VR&S show that this does not work for sentences with a disjunction like P(a) ∨ P(b), because to be competent on the extension of P would entail knowing whether *a* is in the extension of P and whether *b* is. That means the speaker is either in epistemic state s1, s2a or s2c of Figure 1. Eps_1 would then pick s2a and s2c, as they are more minimal than s1. However, if the speaker is in any of those states, she would be withholding information by uttering P(a) ∨ P(b), and violating the maxim of Quantity. This is similar to the reasoning in section 2 why step (vi) and (vii) of the nonce derivation cannot be taken with the disjunct-alternatives, and similar

[5] Actually, as eps_1 also picks out an epistemic state in which P(a)∧P(b) is the case (s3b), it also catches the other paraphrase of the weak SI (*the speaker does not know the truth value of* P(a)∧P(b).)

to Sauerland's reasoning why the secondary implicatures of the disjunct-alternatives are blocked.

Therefore, VR&S take a step similar to the one by Sauerland. They argue we should only add competence as far as it is consistent with the speaker obeying the maxims of Quality and Quantity. As Quality and Quantity were modeled by eps_1, the maximalization of competence should apply after application of eps_1.

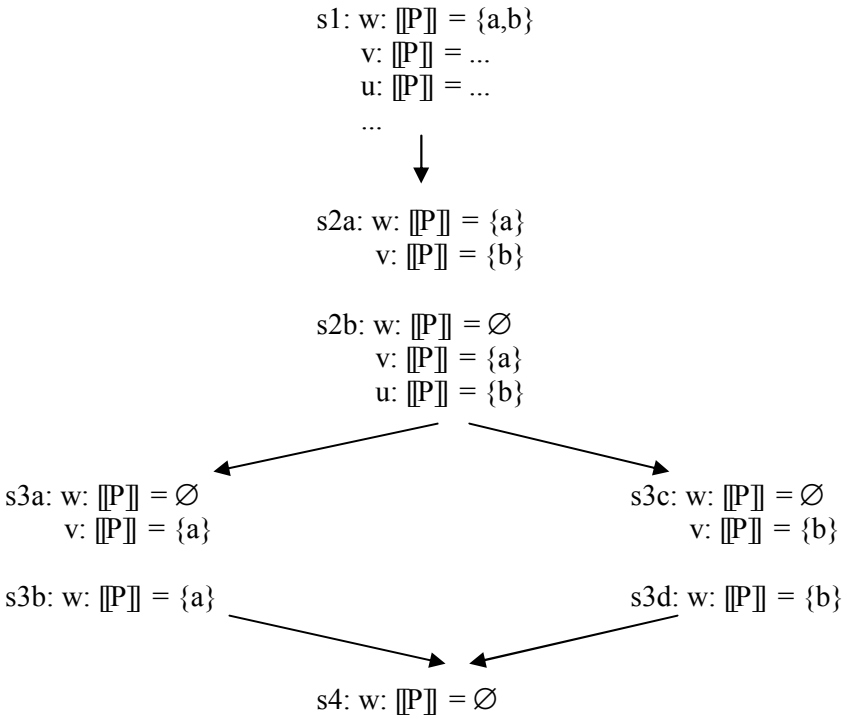
To formalize adding competence, VR&S introduce another exhaustivity operator, eps_2, which makes use of a different ordering $\leq_{\Diamond}$, given in (18):

(18) $s_1 = \langle M_1, w_1 \rangle \in S$:
$s_1 \leq_{\Diamond} s_2$ iff $\forall v_1 \in R_1[w_1]\ \exists v_2 \in R_2[w_2]$
$V_1(P)(v_1) \subseteq V_2(P)(v_2)$

$s1 \cong_{\Diamond} s2$ iff $s_1 \leq_{\Diamond} s_2$ and $s_2 \leq_{\Diamond} s_1$

This formula is very similar to the one of $\leq_{\Box}$, but this time it orders the states on *negative* knowledge: of how many individuals in the domain the speaker knows that they are *not* in the extension of P. This makes sense if we consider that eps_2 will apply to the output of eps_1. Eps_1 picked out states which were similar in positive knowledge about the extension of P. So if we want to pick from those states the one in which the speaker knows most, we have to maximize negative knowledge about the extension of P. The order is formalized by saying that *s* is more minimal than *s'* if for every epistemic possibility in *s*, there is an epistemic possibility in *s'*, in which the extension of P is a superset of the one in *s*. Consider Figure 2, in which the order is given for all possible extensions of P in a domain with two individuals, *a* and *b*.

Figure 2: ordering of eps_2 in a domain with two objects



As this ordering is about maximizing negative knowledge, the state in which the speaker knows of the highest number of elements that they are *not* in the extension of P, is the minimal state of this order, in this case s4.[6] In s1, which represents all states which contain an epistemic possibility that *a* and *b* are in the extension of P, the speaker has least negative knowledge, as there are no elements of which she knows that they are not in the extension of P. In s2a-b, the speaker knows that at least one of *a* or *b* is not in the extension of P (just not which one), so they are more minimal than s1. In s3a-b the speaker knows that *b* is not in the extension of P, and in s3c-d the speaker knows that *a* is not in the extension of P. Therefore, these four states are more minimal than s2a-b. Finally, in s4 the speaker knows that both *a* and *b* are not in the extension of P, so this state is more minimal than s3a-d.

Let us consider how eps_2 works with the simple examples. As I said, eps_2 is applied to the output of eps_1, corresponding to adding competence to

[6] Confusingly, in the *minimal* state of the ordering, the speaker has the *most* (negative) knowledge, contrary to eps_1, where in the minimal state the speaker has *least* (positive) knowledge.

the output of a nonce derivation based on the maxims (steps (vi) and (vii) of the nonce derivation). By applying eps_2 we should be able to derive the strong implicatures. Applying eps_1 to P(b) gave us s2c and s2d of Figure 1, which are the same as respectively s3d and s1 of Figure 2. So among these two, eps_2 selects s3d, as it is more minimal than s1 on $\leq_{\Diamond}$. In s3d, only *b* is in the extension of P, so we derive the epistemic strong (exhaustivity) inference that ¬P(a). Let us consider the example with disjunction, as this is relevant for our current purposes. Applying eps_1 to P(a) ∨ P(b) gave us s3a and s3b in Figure 1, which are the same as respectively s2a and s1 in Figure 2. Out of these two, eps_2 picks s2a, as s2a $\leq_{\Diamond}$ s1. In s2a, only *a* or only *b* is in the extension of P, so P(a)∧P(b) is ruled out. This is how the epistemic strong SI (that the speaker knows P(a)∧P(b) is false) is derived on the account of VR&S.

4.2.2 Van Rooij & Schulz's model and the Speaker Expertise Paradox

Similar to exh^{GS} of Van Rooij (2002), the two exhaustivity operators of VR&S (eps_1 and eps_2) exhaustify the extension of the question predicate in the answer which explains why SIs arise more readily in the focus part of the sentence. However, the proposal of Van Rooij & Schulz also cannot account for the Speaker Expertise Paradox. Why would an addressee assume the speaker has maximal knowledge as to what is *not* in the extension of P (eps_2), after she first assumed the speaker knows as little as possible about what is in the extension of P (eps_1). In other words, why would the addressee maximize competence of the speaker and apply eps_2, after eps_1 has provided the Clausal Implicatures that the speaker does not know the truth value of the disjuncts nor of *A and B*? Again, the only way to avoid this problem is to assume eps_2 (the assumption of maximal knowledge about what is *not* in the extension of P) is default. This could account for the data discussed so far: as eps_1 and eps_2 only apply to the focus of the sentence, it explains the focus-sensitivity of SIs. Contrary to a default version of the account of Sauerland, this can also explain the processing findings: the delay on the scalar region in the focus condition might be due to application of eps_1 and eps_2. However, this would mean that Clausal Implicatures and Scalar Implicatures can never occur independently. Although this is not contradicted by the data of Experiment 4 (the difference between the IgnI-rates and the SI-rates is in the same range as the difference between the baseline condition 4), this prediction might be too strong. In the next chapter I will consider another prediction made by the (original version of the) theory of Van Rooij & Schulz which does not seem to be supported by the data presented there. I return to the option of default exhaustivity in chapter 10.

Summing up, we have seen that the two-step Gricean accounts of Sauerland (2004) and Van Rooij & Schulz (2004) provide elegant solutions for why the strong SI does not lead to exclusion of the disjuncts. However, I argue both proposals can only account for the Speaker Expertise Paradox if the additional assumption is made that the second step is made by default.

Another way around the Speaker Expertise Paradox is abandoning the view that SIs come about by Gricean reasoning. This approach was taken (for independent reasons) by Chierchia (2004, 2006), Chierchia et al. (2008) and by Fox (2007). These theories claim Clausal Implicatures come about by Gricean reasoning, but SIs are derived through application of an exhaustivity operator in the grammar. I will turn to these theories now.

5. Non-Gricean accounts: Grammatical exhaustification

In this section I present two accounts that claim SIs are not derived by Gricean reasoning, but by application of a covert exhaustivity operator in the grammar. I discuss the theory of Fox (2007) first, as it is closer to the Gricean account in the sense that it assumes an exhaustivity operator which acts on the whole sentence, similarly to Van Rooij & Schulz. However, Fox claims this operator is part of the syntactic/semantic system instead of a pragmatic operation. Chierchia (2004, 2006) and Chierchia, Fox & Spector (2008) propose SIs are derived by an exhaustivity operator which can be applied locally during compositional semantics.

5.1 Fox (2007): grammatical exhaustification at the sentence level

5.1.1 The mechanism

Fox (2007) points to the similarity between overt *only* and SIs, and proposes the *only* implicature generalization (OIG):

(19) The *only* implicature generalization (OIG):
A sentence, *S*, as a default, licenses the inference/implicature that (the speaker believes) *onlyS′*, where *S′* is a modification of *S* with focus on scalar items.[7]

[7] Notice that by *focus* Fox means stress / pitch accent, so not information focus, which is the meaning of the word *focus* I use in this work.

Fox proposes to assume a covert exhaustivity operator (*exh*), similar to *only*, but for which the requirement that the prejacent (the sentence without *only*) is true is part of the assertion, and not a presupposition. Like the exhaustivity operator of Van Rooij (2002), Fox's exhaustivity operator is based on the operator proposed by Groenendijk & Stokhof (1984). Fox defines *exh* as in (20):

(20) $[\![\text{Exh}]\!](A_{<<s,t>,t>})(p_{<s,t>})(w) \Leftrightarrow p(w) \,\&\, \forall q \in NW(p,A): \neg q(w)$

(20) says that the proposition is true and all non-weaker (NW) alternatives are false. The SIs is derived by applying *exh* to the sentence in the grammar.

However, Fox wants to hold on to a non-restrictive set of alternatives, which includes the disjuncts themselves. This is problematic, as the disjuncts are also non-weaker alternatives. So (20) would also rule them out, which would lead to a contradiction. Fox therefore adopts Sauerland's view that only those alternatives can be excluded that do not lead to inclusion of other alternatives. He calls these *innocently excludible* (I-E) alternatives. Fox incorporates this constraint in the exhaustivity operator in (21):

(21) $[\![\text{Exh}]\!](A_{<<s,t>,t>})(p_{<s,t>})(w) \Leftrightarrow p(w) \,\&\, \forall q \in NW(p,A)$: [q is *innocently excludable* given $A \rightarrow \neg q(w)$]

where q is *innocently excludable* given A if: $\neg\exists q' \in NW(p,A)$ $[p \wedge \neg q \Rightarrow q']$

So any non-weaker alternative proposition q is excluded ($\neg q(w)$) if it does not lead to inclusion of another non-weaker proposition q'. For a disjunctive proposition like $A \vee B$, this replicates Sauerland's results. As neither of the disjuncts can be excluded without including the other, they are both not innocently excludable. The only alternative that can be innocently excluded is $A \wedge B$, giving rise to the SI.

However, Fox notices that one problem remains for his (and Sauerland's) system of innocent exclusion: the sum of all exclusions can lead to a contradiction. For instance, all individuals in a domain are non-weaker alternatives to the answer *somebody*, and each of them can be innocently excluded. But excluding all of them leads to the answer *nobody*, which contradicts the assertion, a problem also noticed by Van Rooij 2002. Therefore, Fox changes the notion of *innocently excludible*: to see if a proposition is innocently excludible, we have to look at the set of possible maximal exclusions which are consistent with the assertion. Only the alternatives that are in each of these sets, are innocently excludible. In other

words, only the alternatives that are in the intersection of all the sets of maximal exclusions, can be excluded. This is formalized in (22):

(22) $[\![\text{Exh}]\!](A_{<<s,t>,t>})(p_{<s,t>})(w) \Leftrightarrow p(w) \,\&\, \forall q \in \text{I-E}(p,A) \rightarrow \neg q(w)$

I-E(p,A) = $\cap\{A' \subseteq A$: A' is a maximal set in A such that $A'^{\neg} \cup \{p\}$ is consistent$\}$ $A'^{\neg} = \{\neg p: p \in A\}$

For a sentence like $A \vee B$ and the Sauerland scale in (13) above, there are two maximal exclusions that are consistent with the assertion: $\{A, A \wedge B\}$ and $\{B, A \wedge B\}$. The intersection of these two sets is $\{A \wedge B\}$, so the alternative $A \wedge B$ can be innocently excluded, which gives us the SI.[8] So the approach of Fox derives SIs by a grammatical exhaustivity operator which acts on a scale like Sauerland's, where the disjuncts themselves are also alternatives.

The crucial difference between Fox' account and the account of Van Rooij & Schulz (2004) is that the exhaustivity operator derives the SIs in the grammar, and the Ignorance Implicatures are derived by Gricean reasoning based on the output of the grammar. So the order of calculating SIs and IgnIs is reversed compared to Van Rooij & Schulz, and the SI does no longer depend on Gricean reasoning.

5.1.2 Fox's model and the Speaker Expertise Paradox

Fox's account avoids the Speaker Expertise Paradox as he claims the SI is derived in the grammar, not by Gricean reasoning. However, the question remains when this grammatical exhaustification procedure is activated. Fox claims this happens when the hearer considers the Ignorance Inferences that she derived by Gricean reasoning, to be implausible. He illustrates with (23):

(23) I ate the cake or the ice cream.

If the hearer parses (23) without an exhaustivity operator, Gricean reasoning gives rise to the Ignorance Implicatures that the speaker does not know the truth value of *I ate the cake*, *I ate the ice cream* and *I ate the cake and the ice*

[8] The problem with an answer like *somebody* is also solved. Say we are in a domain with three individuals, a, b and c. Then every maximum exclusion will contain all the alternatives (so $\{a\}$, $\{b\}$, $\{c\}$, $\{a \wedge b\}$, $\{b \wedge c\}$, $\{a \wedge c\}$, $\{a \wedge b \wedge c\}$), but one of $\{a\}$, $\{b\}$ and $\{c\}$. The intersection of these exclusions is $\{\{a \wedge b\}, \{b \wedge c\}, \{a \wedge c\}, \{a \wedge b \wedge c\}\}$. So the possibilities with more than one person are excluded, leaving $\{\{a\},\{b\},\{c\}\}$ as possibilities, which is correct for the answer *somebody*.

cream. The hearer might judge this implausible and therefore try another parse in which the exhaustivity operator is applied. As explained above, only the alternative *I ate the cake and the ice cream* can be innocently excluded, and the participant ends up with the exclusive-*or* meaning.

However, by introducing this dependence on implausible Ignorance Implicatures, Fox actually turns the application of exhaustivity into a pragmatic device. If a re-parse of the sentence in which the exhaustivity operator is applied depends on the Ignorance Implicatures being considered implausible by the hearer, it depends on pragmatic reasoning of the hearer. This re-introduces the Speaker Expertise paradox. In (23) it might be implausible that the speaker does not know the truth value of *A and B*, as the sentence is about what she herself did. However, if we take the examples considered so far, the SEP re-arises. Consider (24):

(24) Laura found a crab or a mussel.

On Fox's view, Gricean reasoning here gives rise to the Ignorance Implicatures that the speaker does not know the truth value of *Laura found a crab*, *Laura found a mussel* and *Laura found a crab and a mussel*. Why would the addressee find these inferences implausible and go for a second parse in which the exhaustivity operator is applied? Similar to what I argued for the accounts of Sauerland and Van Rooij & Schulz, we have to assume that this assumption is default to be able to derive the exclusive-*or* meaning. It seems that deriving the SI in the grammar does not help if this derivation depends on Gricean reasoning about an initial parse without SI. I will now discuss an exhaustivity approach which allows application of the exhaustivity operator during compositional semantics, the approach by Chierchia (2004, 2006).

5.2 Chierchia (2004, 2006) and Chierchia, Fox and Spector (2008): local grammatical exhaustification

In chapter 2 I already briefly introduced the approach of Chierchia (2004, 2006) because of its link to focus semantics. Chierchia rejects the Gricean view of SIs because he claims it fails to account for certain cases where the SI arises embedded under another operator. The Gricean view is a pragmatic account that applies to a semantic representation of a whole sentence (the output of compositional semantics). Chierchia shows that in some cases, SIs appear to be calculated in embedded positions, instead of on the whole sentence meaning. Therefore, he proposes a radically different account of SIs, which allows SIs to be calculated during compositional semantics (see

my discussion of the globalist-localist debate in section 2 of chapter 5). I will not go into the globalist-localist issue, but I will discuss Chierchia's approach as it derives SIs in the grammar and therefore it possibly escapes the Speaker Expertise Paradox.

5.2.1 The mechanism

The system Chierchia proposes for SIs is similar to alternative semantics for focus. He states that every expression α has, next to its plain meaning $[\![\alpha]\!]$, a 'scalar' or 'strengthened' value $[\![\alpha]\!]^S$. This strengthened value comes about by considering the set of alternatives to α, $[\![\alpha]\!]^{ALT}$, which is defined as (25):

(25) $[\![\alpha]\!]^{ALT} = \{ \quad \{a_1,..., a_n\}$, if α is part of a scale $<a_1,... ,a_n>$
$\{a\}$ otherwise

Two things follow from this definition of $[\![\alpha]\!]^{ALT}$. First, the alternatives are lexically constrained, i.e. they rely on a lexically given scale (e.g. a Horn scale). Secondly, if an utterance does not contain a scalar term, $[\![\alpha]\!]^{ALT}$ is the set containing only the plain meaning. Chierchia (2004) proposes the following rule (26), which introduces the implicatures through exhaustification.

(26) If φ is a scope site (of type t), then $[\![\varphi]\!]^S = [\![\varphi]\!]^S \wedge \neg S([\![\varphi]\!]^{ALT})$

In this formula, $S([\![\varphi]\!]^{ALT})$ is the member of the alternative set immediately stronger than φ (so the expression with the scalar term replaced by a scalar term that is one step higher on the scale). It is easy to check that by rule (26), if φ is *John has A or B*, $[\![\varphi]\!]^S$ is *John has A or B and it is not the case that John has A and B*.

In Chierchia (2006), the exhaustivity operator O is introduced, to replace the rule in (26) above. Chierchia defines O as in (27), where C is the set of alternatives to an expression φ, $[\![\varphi]\!]^{ALT}$.

(27) $O_C[q] = q \wedge \forall p[((p \in C) \wedge p) \rightarrow (q \subseteq p)]$

The strengthened meaning of an expression φ, represented by $[\![\varphi]\!]_S$, is the result of applying O to the plain meaning and its alternatives, see (28):

(28) $[\![\varphi]\!]_S = O_C [\![\varphi]\!]$, where $C= [\![\varphi]\!]^{ALT}$

The working of O is similar to the mechanism in (25) and (26), but it incorporates the *Strength Condition* of Chierchia (2004), which says the strong value cannot become weaker than the plain value. The operator O applies to a meaning q and says that every alternative p from the set of alternatives C is entailed by q (and therefore the alternatives that are not entailed by q are false). As the scale for *or* that is assumed by Chierchia does not contain the disjuncts, for a simple disjunction A∨B, the alternative set C is the set {A∨B, A∧B}. Therefore, O(A∨B) = (A∨B) ∧ ¬(A∧B), which gives us the SI. In Chierchia, Fox and Spector (2008), this model is elaborated on.

5.2.2 Chierchia (and colleagues)' model and the Speaker Expertise Paradox

Similar to Fox's (2007) account, the account of Chierchia and colleagues avoids the SEP as the SI is not derived by Gricean reasoning but by an exhaustivity operator in the grammar. However, Chierchia (2004, 2006) and Chierchia et al. (2008) do not provide a clear criteria for the circumstances in which exhaustivity is applied. They mostly focus on the interaction of the exhaustivity operator with structural factors such as monotonicity. However, they do point out that the exhaustivity operator is *optional*:

> 'This optionality is also captured by our grammatical mechanism. Given what we've said up to now, there is nothing that forces the presence of the operator O in a sentence containing a scalar item. Optionality is thus predicted, and one can capture the correlation with various contextual considerations, under the standard assumption (discussed in the very beginning of this paper) that such considerations enter into the choice between competing representations (those that contain the operator and those that do not).' (Chierchia, Fox & Spector 2008, p. 28)

Similar to the objection against Fox (2007), on this view the question arises why hearers would apply the exhaustivity operator (or pick the representation in which it is applied) in an episodic UE-sentence like *Laura found a crab or a mussel*, which also gives rise to Ignorance Implicatures. Therefore, it seems the problem of the Speaker Expertise Paradox is not solved. However, Chierchia et al. go on:

> 'However, the assumption that the operator O is optional is not a *necessary* assumption. One might assume instead that there is an optional process that activates the alternatives of a scalar item, but that

> once alternatives are active, SIs are obligatory (see Chierchia 2006 for an implementation).' (Chierchia, Fox & Spector 2008, p. 28, their italics)

The implementation of Chierchia (2006) they refer to is that scalar terms could come in two versions. One version in which the alternatives are active and one in which they are not. This is the relevant passage of Chierchia (2006):

> 'A characteristic of scalar inferences is that they can be suspended. If we assume that scalar terms activate alternatives by default and that alternatives must be factored into meaning, how is suspension of implicatures possible? We must assume that the default activation of alternatives can be, in turn, suspended. A simple way to achieve suspension is to assume that each scalar item comes in two variants (say, thanks to an abstract morphological feature [±σ], where σ is a mnemonic for strong): *many*$_{[\pm\sigma]}$, *or*$_{[\pm\sigma]}$, and so on; a [+σ] item (e.g., *many*$_{[+\sigma]}$) has active alternatives and must lead to enrichment, while a [−σ] item (e.g., *many*$_{[-\sigma]}$) has no active alternatives and cannot lead to enrichment. Speakers choose the feature setting that fits the context best.' (Chierchia 2006, p. 547)

This only seems to be moving the problem from why hearers would apply the exhaustivity operator in episodic UE-sentences to why participants would pick the strong *or*$_{[+\sigma]}$ in episodic UE-sentences. However, Chierchia et al. go on to point out that some items seem to only have a strong variant, i.e. they obligatorily trigger the alternatives and therefore obligatorily trigger exhaustification. They illustrate this possibility of obligatory SIs with an idea of Spector (2007), who considered a contrast in the interpretation of the plural morpheme, see (29) and (30):

(29) John read books.

(30) a. John didn't read books.
b. I don't think that John read books.

In (29) *books* seems to mean *more than one book*, which is illustrated by the infelicity of (31):

(31) # John read books. In fact, he read exactly one book.

However, both sentences in (30) would be judged false if John read one book. Therefore, it seems in (30) *books* means *at least one book*. Spector proposed we can account for this discrepancy by assuming that the *at least one* meaning is basic and the *more than one* meaning comes about by an SI. The exact implementation of this idea is quite complex, so I skip it here and refer the interested reader to Spector (2007)[9] What is interesting for our current purposes is that Chierchia and colleagues claim that the plural morpheme might be a term which obligatory activates its alternatives and therefore leads to obligatory exhaustification: [10]

> 'Specifically, under the grammatical theory, one could claim that for plural morphology, in contrast to standard scalar terms, alternatives are automatically activated. Once alternatives are active, they must be associated with the operator O, and – to the extent that some alternatives are stronger than the basic sentence – implicatures are obligatory.' (Chierchia et al. 2008, p. 30)

The fact that in (30) the SI did not arise despite obligatory exhaustification is that in DE-environments such as (30) exhaustification (applied at the top-level) does not lead to a stronger meaning and hence the operation of O is vacuous.

---

[9] Chierchia et al. 2008 also skip it so it is not crucial for their argumentation.

[10] However, I think that the plurality of *books* in (29) is just as focus-sensitive as the SI of *or*. Consider (i):

(i) [context in which John read one book, Jack read a magazine and Pete read nothing.]
Q: "Who read books?"
A: "$\underline{\text{John}}_F$ read books."

My intuition is that participants will be just as likely to judge the answer 'true' as in the TVJT-experiments with *or*.

This gives rise to interesting SPR conditions, which I give here as a suggestion for further research:

(i) <u>focus condition</u>
A: "What did John read on vacation?"
B: "John read books on vacation."
John said he read it/them in the car.

(iii) <u>non-focus condition</u>
A: "Who read books on vacation?"
B: "John read books on vacation."
John said later he read it/them in the car.

The critical measure would be a delay on *it* in (i).

So ultimately, we could propose *or* is like the plural morpheme in that it always has active alternatives and therefore leads to obligatory exhaustification. This could explain why the exclusive reading of *or* is observed even though no speaker expertise is assumed. However, for episodic UE-sentences this reduces the account of Chierchia et al. to a defaultist account, and the focus-sensitivity becomes hard to explain. In the non-focus conditions I also used UE episodic sentences, so if exhaustification is obligatory for *or*, it should be calculated there too. I conclude that the approach of Chierchia and colleagues also cannot account for the Speaker Expertise Paradox and the focus-data at the same time. Nevertheless, it is an interesting suggestion that *or* is more like the plural morpheme than like other scalar terms. I discuss this option further as a direction for future research in chapter 10.

6. Summary and conclusions

In this chapter I looked into a well-known problem for the SI-view of *or*, which I dubbed the Speaker Expertise Paradox. As *or* also triggers Clausal / Ignorance Implicatures, on the simple Gricean view of SIs the hearer must have made a paradoxical assumption about the epistemic state of the speaker. The hearer must have taken the speaker to be uninformed about the truth value of the disjuncts, but informed about the truth value of *A and B*. Additional data of Experiment 4 showed that IgnIs and SIs indeed co-occurred. Participants often calculated SIs even though they did not consider the speaker to be completely informed. Therefore, I conclude the Speaker Expertise Paradox is a serious problem for the Gricean view of the SI of *or*.

I discussed two approaches in the literature that tried to save the Gricean account by proposing a two-step theory of deriving Clausal Implicatures and SIs: the accounts of Sauerland (2004) and Van Rooij & Schulz (2004). Although these proposals show how calculation of the strong SI does not necessarily lead to exclusion of the disjuncts too, I argue they do not solve the Speaker Expertise Paradox. It is unlikely hearers would take the second step, which derives the strong SIs, after the weak Clausal Implicatures (and the weak SI) have been derived in the first step.

I also considered two accounts which abandon the view that SIs come about by Gricean reasoning and for that reason might avoid the problem. Fox (2007), Chierchia (2004, 2006) and Chierchia et al. (2008)) claim SIs are derived through application of an exhaustivity operator in the grammar. However, Fox (2007) re-introduces the Speaker Expertise Paradox by having a re-parse of the sentence with application of the exhaustivity operator depend on pragmatic reasoning of the hearer. Chierchia et al. first claim their

exhaustivity operator is optional, which gives rise to the SEP. Then they claim selecting the scalar term with active alternatives is optional, which only seems to be moving the problem. Finally, they point out some items have only a strong version which gives rise to obligatory exhaustification. If we assume *or* is one of those special items we do solve the Speaker Expertise Paradox, but we cannot explain the effects of focus observed in the previous chapters.

I conclude that (in their current form) neither the simple Gricean approach, nor the extensions of Sauerland (2004) and Van Rooij & Schulz (2004), nor the grammatical exhaustivity accounts of Fox (2007), Chierchia (2004, 2006) and Chierchia et al. (2008) can solve the Speaker Expertise Paradox and account for the focus-sensitivity at the same time. Only a version of the exhaustification account of Van Rooij & Schulz in which the second step is default can account for both. This raises considerable doubt about whether the view that the exclusive reading of *or* comes about by an SI, is correct.

In the next chapter I follow up on the accounts introduced in this chapter. As three of them explicitly link SIs to exhaustivity, I consider the relation between SIs and exhaustivity, which was tested in Experiments 5-7. The results again raise doubts about the derivation of exclusive-*or* by an SI.

CHAPTER 9: THE RELATION BETWEEN THE SI OF *OR* AND EXHAUSTIVITY

1. Introduction

In chapters 3-7 I tested the prediction that SIs are sensitive to focus. One of the theories which made this prediction was the theory of Van Rooij (2002). He claims SIs arise through a more general procedure of *exhaustification of answers*. An example of an exhaustivity inference is that the answer in (1) is usually taken to mean (2):

(1) Q: Who brought a laptop to class?
A: John brought a laptop to class.

(2) John brought a laptop to class and nobody else did.

The part of (2) that *nobody else did* is the exhaustivity inference. The link with SIs is straightforward, as both seem to be the result of the addressee assuming that the speaker is trying to be as informative as possible (Grice's maxim of Quantity). In (1), if someone else (in the relevant domain) also brought a laptop to class, the speaker should have said so. From the fact that she did not, we can conclude that no one else did. This is why according to Van Rooij (2002) (and the more recent versions of the proposal in Van Rooij & Schulz (2004) and Schulz & Van Rooij (2006)), we can consider exhaustification as a way of making Gricean reasoning explicit.

According to Van Rooij and many other authors (e.g. Groenendijk & Stokhof 1984), exhaustification of answers is sensitive to focus. Exhaustivity usually affects the part of the sentence that provides the answer, not the part in which the question predicate is repeated.[1] As the part that provides the answer is the focus, exhaustivity only affects the focus. For instance, the answer in (1) does not give rise to the inference in (3):

(3) John did not bring anything else to class.

[1] This description is sloppy as I want to get the intuition across. Actually, what is exhaustified is the extension of the question predicate in the answer. This results in a reduction of the set of sets denoted by the material in the focus. See section 4.2 of chapter 2 for a more elaborate and more accurate description.

In other words, as *John* is the focus part of (1A), we do make the exhaustivity inference from *John* to *only John*, but as *a laptop* is background, we do not draw the inference from *a laptop* to *only a laptop*.

Van Rooij argues that SIs come about by the same mechanism (the application of an exhaustivity operator) and therefore they should only arise if the scalar term is in the focus. In Experiments 1-9 I tested question answer pairs like (4) and (5):

(4) Q: What did Harry bring?
A: Harry brought bread or chips$_F$.

(5) Q: Who brought bread or chips?
A: Harry$_F$ brought bread or chips.

We observed that the SI of *or* (*Harry did not bring bread and chips*) was indeed calculated more often in (4) than in (5), in line with Van Rooij's predictions.

In this chapter I consider the question whether the observed focus-sensitivity of SIs is indeed due to SIs being a form of exhaustivity. I do this by considering additional data from Experiments 5-7, which next to items that tested SIs contained items that tested exhaustivity. In section 2 I discuss the predictions made by Van Rooij's exhaustivity approach of SIs. There, I also show that these predictions do not just follow from the account of Van Rooij (2002) / Van Rooij & Schulz (2004), but also from the accounts which assume a grammatical exhaustivity operator: Fox (2007) and Chierchia et al. (2008), introduced in the previous chapter. In sections 3 and 4 I present the additional data from Experiments 5-7. Section 5 contains the general discussion and conclusions. Although (as I already pointed out in section 7.2 of chapter 4) the data might be affected by the experimental question overruling the critical manipulations, the results raise doubts about whether SIs and exhaustivity arise by the same mechanism.

2. The prediction of co-occurrence of SIs and external exhaustivity

2.1 The prediction

The view that SIs and exhaustivity inferences arise by the same exhaustivity operator gives rise to an interesting prediction, which is best explained by an example. Consider (6):

(6) Q: What did Bart catch?
A: Bart caught a spider or a snake.

As *or* is in the focus part of the sentence the exhaustivity operator applies to it, and the SI in (7) is predicted to arise:

(7) Bart did not catch a spider and a snake.

However, next to the SI, the application of the exhaustivity operator should also give us the inference in (8), similar to what happened in (1) and (2) above:

(8) Bart did not catch anything else (besides a spider or a snake).

To be able to refer to the two effects of exhaustification, I will refer to the type of exhaustification in (8) (the exclusion of non-scalar alternatives) as *external exhaustification*. In Experiments 5-7, I tested both the SI in (7) as the external exhaustification inference in (8).

The crucial prediction of the exhaustivity view of Van Rooij is that the two inferences should always co-occur. Both are derived by a single application of the exhaustivity operator. Therefore, if one of them is observed that means the exhaustivity operator was applied, and the other one should also be observed. Correspondingly, as both inferences are the result of the same process, they should be triggered to the same extent in the same environments. Experiments 5-7 contained three test conditions in which focus and alternative relevance were manipulated. If SIs and external exhaustifications are the result of the same mechanism, they should be sensitive to the manipulations of focus and alternative relevance to the same degree. Therefore, in all three conditions similar rates of SIs and external exhaustification inferences should be observed. This prediction was tested in Experiments 5-7.

I argue that not only the theory of Van Rooij (2002) and Van Rooij & Schulz (2004) make this prediction, but also the accounts of Fox (2007) and Chierchia et al. (2008). In the next subsections I show how this prediction is derived from the mechanics of their proposals. The easiest way to check the predictions of necessary co-occurrence of the SI of *or* and external exhaustivity, is to assume a domain with three objects, *a*, *b* and *c*, and consider whether the SI of *a or b* (*not a and b*) necessarily co-occurs with the external exhaustivity inference *not c*. Consider for instance (9):

(9) [domain with three objects: a, b, c]
Q: What does John have?
A: John has a or b.
SI: John does not have a and b.
external exh: John does not have c.

First, for completeness I show how this prediction is derived on the account of Van Rooij & Schulz (2004). Then I show how it is derived by Fox (2004) (and with minimal additional assumptions by Sauerland (2004)). Finally, I show how it is derived on the view of Chierchia et al. (2008). This requires an addition to their theory made by Fox & Spector (2008). In the following subsections, I take the mechanisms as presented in sections 4 and 5 of the previous chapter as a starting point.

2.2 Co-occurrence of the SI of *or* and exhaustivity on the account of Van Rooij & Schulz (2004)

As I pointed out above, Van Rooij (2002) and Van Rooij & Schulz (2004) (VR&S) explicitly make the claim that SIs come about by the same mechanism as external exhaustification. Let us nevertheless consider how the prediction of co-occurrence of SIs and external exhaustivity is derived from the exhaustivity mechanism of VR&S. I focus on the example given in (9) above.

I will not attempt to give the full ordering eps_1 gives rise to in a domain with three individuals (a, b and c), as the number of possible states is quite large, but I will limit myself to finding the minimal states among the states in which the speaker knows a∨b, as this is relevant for our example. In (9), the answer is a∨b. We first have to select the states in which the speaker knows a∨b. In order for the speaker to know a∨b in a state, every epistemic possibility in that state has to contain at least either *a* or *b*. The state *s* in (10) is an example of this quite large set of states:

(10) s: v: [[P]] = {a}
w: [[P]] = {a,c}
u: [[P]] = {b,c}

Now we have to determine which of these epistemic states are the minimal ones on $\leq_{\square}$. The states containing both the epistemic possibility {a} and the epistemic possibility {b} will be the minimal states, as for every superset of {a} or {b}, these states contain a subset, which is the requirement for a state to be more minimal on $\leq_{\square}$. This makes sense, because in these states the

speaker does not know whether *a* is in the extension of P, nor whether *b* is, nor whether *c* is. The requirement also says that the inverse should not hold (otherwise the two states are equally minimal). This part of the requirement is also satisfied, as all the states that do not contain both {a} and {b}, will not have a subset of {a} or will not have a subset of {b}. The minimal states picked out by eps_1 therefore are the ones which contain both the atomic possibilities {a} and {b}. See (11), where the three dots represent additional possibilities with at least a or b:

(11) v: [[P]] = {a}
w: [[P]] = {b}
...

In (12) below I list these states, by (for reasons of space) putting the epistemic possibilities next to each other instead of under each other. For clarity, s1 in (12) is (11) above without the three dots. This notation is different from the one of Van Rooij & Schulz, who also list epistemic states as sets, which is somewhat confusing.

(12) minimal states picked out by eps_1 in a domain of three individuals (a,b,c), where the speaker knows that a∨b:

s1: {a},{b}
s2: {a},{b},{a,b}
s3: {a},{b},{a,c}
s4: {a},{b},{b,c}
s5: {a},{b},{a,b,c}
s6: {a},{b},{a,b},{a,c}
s7: {a},{b},{a,b},{b,c}
s8: {a},{b},{a,b},{a,b,c}
s9: {a},{b},{a,c},{b,c}
s10: {a},{b},{a,c},{a,b,c}
s11: {a},{b},{b,c},{a,b,c}
s12: {a},{b},{a,b},{a,c},{b,c}
s13: {a},{b},{a,b},{a,c},{b,c},{a,b,c}

These states are all equally minimal, as they all contain {a} and {b}, and all additional possibilities always contain *a* or *b*, so they always contain a subset of every additional possibility in every other state.

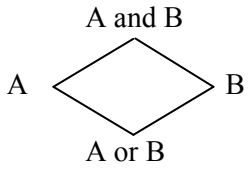
Now that eps_1 has picked the states in (12), it is time to add competence and see which states are selected by eps_2. Using the requirement on $\leq_\Diamond$, that one state is more minimal than the other if for every possibility of the first

one, there is a superset in the latter (and the inverse does not hold), it is clear that eps_2 will pick out s1, which only contains the possibilities {a} and {b}. Similar to the example with a domain of two elements, this means states containing {a,b} are ruled out, giving rise to the strong SI reading. However, also states containing possibilities {a,c}, {b,c} and {a, b, c} are ruled out. This means that from *a or b*, if we assume the speaker is competent, we take her to know that there is no third element (c) in the extension of the question predicate P. So as expected, the account of VR&S predicts co-occurrence of the SI of *or* and external exhaustivity.

2.3 Co-occurrence of the SI of *or* and exhaustivity on the account of Fox (2007) (and Sauerland (2004))

Fox (2007) adopts the scale of Sauerland (2004), in which the disjuncts are also alternatives, repeated here in (13). As Fox also adopts Sauerland's system of innocent exclusion, I collapse their views here.[2]

(13)



The reason for adopting the Sauerland scale is that Fox wants to hold on to what he calls the *Basic Maxim of Quantity* (B-MQ), which states *any* utterance that is relevant and more informative should be preferred over one that is less informative. What does the alternative set look like for the example in (9) on B-MQ? I first repeat the example:

(9) [domain with three objects: a, b, c]
Q: What does John have?
A: John has a or b.
SI: John does not have a and b.
external exh: John does not have c.

Fox gives a similar example, where he argues the alternatives are determined by focus. For the question *Who did Fred talk to?* he proposes the set of

[2] However, Sauerland does not mention alternatives other than the ones in (13). Nevertheless, it is interesting to explore whether his approach can be extended.

alternatives for the answer is the following Hamblin-denotation of the question:

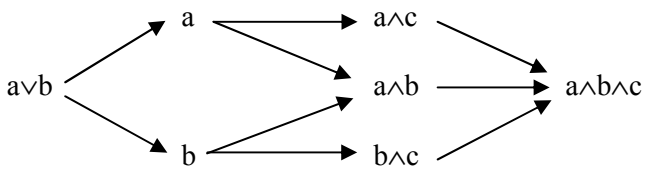
(14) {that Fred talked to x: x is a person or a set of people}

Adopting this system, the alternatives for (9) is the set in (15):

(15) {that John has x: x is an object or a set of objects in our domain}

The set in (15) corresponds to {{a},{b}, {c}, {a,b}, {a,c}, {b,c}, {a,b,c}}, (where the letters represent the propositions that John has those objects). Now, we can make a version of the Sauerland and Fox scale in (13), with these alternatives and the assertion a∨b. This results in (16), where the arrows point to the stronger alternatives.[3] [4]

(16)



If we apply Sauerland's system to this scale, we get the following Primary Implicatures:

(17) a. The speaker does not have the positive belief that (a∧b∧c).
b. The speaker does not have the positive belief that (a∧b)
c. The speaker does not have the positive belief that (a∧c).
d. The speaker does not have the positive belief that (b∧c).
e. The speaker does not have the positive belief that a.
f. The speaker does not have the positive belief that b.

As in the case with two objects, the epistemic step can be taken for (17b), giving rise to the strong SI, while it cannot be taken for (17e) and (17f), yielding the weak Clausal Implicatures. However, we can also take the epistemic step for all the alternatives that contain *c*: (17a), (17c) and (17d), resulting in the following set of secondary implicatures:

[3] For notational convenience, the scale is represented horizontally instead of vertically as in (13).

[4] The alternative c is left out, as it has no entailment relation with the assertion a∨b.

(18) a. The speaker knows (a∧b∧c) is false.
b. The speaker knows (a∧b) is false.
c. The speaker knows (a∧c) is false.
d. The speaker knows (b∧c) is false.

So if we apply Sauerland's system to a scale like (16), we predict the SI and external exhaustification (that John does not have *c*) to co-occur. If we can take the epistemic step for the a∧b alternative, we can also take it for every alternative containing c.

Fox extended Sauerland's system by not just looking at which alternatives are innocently excludable by themselves, but by taking the intersection of the possible maximal exclusions. For the scale in (16), there are two maximal exclusions: the set containing every alternative except the assertion (a∨b) and a, and the set containing every alternative except the assertion and b. The intersection of these maximal exclusions therefore is the set containing every alternative except the assertion, the alternative *a* and the alternative *b*. Crucially, this set contains both the alternatives in which a∧b is the case (a∧b, a∧b∧c), as all alternatives which contain *c* (a∧c, b∧c and a∧b∧c). Therefore, these can be innocently excluded, leading to the inferences that John does not have *a and b* and does not have *c*. So Fox's proposal also predicts co-occurrence of SIs and external exhaustification.

2.4 Co-occurrence of the SI of *or* and exhaustivity on the accounts of Chierchia (2004, 2006) and Fox & Spector (2008)

Chierchia (2004, 2006) assumes that the set of alternatives is lexically constrained, and that the alternative set for *or* only contains *or* and *and*. On that account, the SI of *or* and external exhaustification do not necessarily have to co-occur: they could be two processes with two different alternatives sets. However, from a conceptual point of view this is not very nice, as two separate exhaustification procedures, with two sets of alternatives, would have to be assumed (and possibly also two exhaustivity operators). Also, the inspiration for Chierchia's exhaustivity approach was actually the *similarity* of exhaustification of focus alternatives and scalar alternatives (see chapter 2). Therefore, it seems reasonable to assume that on Chierchia's approach, the scalar alternative set is somehow related to the focus alternative set, and one exhaustivity operator derives both SIs and external exhaustifications.

Fox & Spector (2008) actually make this claim. In an attempt to account for the contrast in (19) and (20) below, they assume that the exhaustivity operator applies to one alternative set C (potentially containing both scalar and focus alternatives), which is derived from the focus alternative set (also

called *focus set* or *focus value*, see chapter 2). They also provide evidence for the prediction of co-occurrence of the SI of *or* and external exhaustivity.

(19) John didn't do the reading OR the homework. He did both.

(20) # John didn't do the reading or the homework. He did both.

Fox & Spector claim the alternative set C is derived from the focus set according to two constraints that they call *Association with Focus (AS)* and *Minimize Focus (MF)*. *AS* is the constraint which Rooth (1992) calls the *Focus Interpretation Principle (FIP)*, which is that the set of alternatives C must be a subset of the focus set. An argument in favor of this constraint (instead of just using the focus set as the alternative set), is that a sentence like (21), taken from Rooth (1992), would otherwise negate all other relations that Mary has to the book 'The Recognitions' (assuming that the focus set is the set of all possible relations between Mary and The Recognitions), while there probably are many relations between the two that are true (e.g. Rooth mentions the relation of being born in the same millennium as the author).

(21) Mary only $[\text{read}]_F$ The Recognitions.

To account for this, Rooth proposes the FIP, which introduces a set of salient alternatives C which is a subset of the focus set. In this case C could be the set {reading The Recognitions, understanding The Recognitions}.

The second constraint on the alternative set C, dubbed *MF* by Fox & Spector, also stems from work on association with focus and is given in (22):[5]

(22) A sentence can't have a focus value F, if it would satisfy AF (= Rooth's FIP) with another focus value F' (derivable by a different distribution of focus marking), and $F' \subset F$.

Now consider (23) and (24), which only differ on placement of stress, which results in different focus structures:

(23) John talked to $[\text{Mary or SUE}]_F$.

[5] Fox & Spector themselves mention Schwarzchild (1999), probably because MF can be considered a version of Schwarzchild's *Avoid F* constraint, which says that speakers should F-mark as little as possible.

(24) John talked to Mary [OR]$_F$ Sue.

(23) is the normal intonational pattern for this sentence, and one of the possible focus structures is that the whole constituent *Mary or Sue* is the focus (e.g. after the question *Who did John talk to?*). This is the type of sentences I have been testing throughout this thesis. Following Rooth (1992), the focus value of a sentence is formed by replacing the focus with other expressions of the same semantic type. In (23), the semantic type is (groups of) individuals, so the focus value of (23) is: {John talked to Mary or Sue, John talked to Mary, John talked to Sue, John talked to Mary and Sue, John talked to Dick,...} In (24), the semantic type is logical connectives, so the focus value for (24) is: {John talked to Mary or Sue, John talked to Mary and Sue}. Now the question is: What is the set of alternatives C in (23) and in (24), that will be ruled out by application of the exhaustivity operator? Fox & Spector point out that for (23), C cannot be just {John talked to Mary and Sue}. Their reasoning is as follows: If we take C = {John talked to Mary and Sue} for (23), we satisfy AF. However, then there is a another focus value F' (namely the focus value of (24)), which would also satisfy AF (because C is also a subset of F'), and which is a subset of the focus value F of (23). Therefore, there has to be at least one alternative in C of (23), that is not in C of (24), e.g. *John talked to Dick*. So Fox & Spector point out that the alternative set C in (23) has to contain *John talked to x*, where *x* is someone else than Mary or Sue.

As the alternatives in C are the ones that are ruled out by the exhaustivity operator, and the exhaustivity operator is responsible for the SI, this predicts that whenever the SI reading is observed, external exhaustivity (that nothing else than *A or B* is the case) should also be observed. Fox & Spector explicitly make this claim:

> 'Consequently, if (53) (here (23)) yields the *"not and"* inference, it must yield an additional exclusive inference that would make it stronger than (52) (here (24)), e.g. ¬D [*John did not talk to Dick*, AZ].' (Fox & Spector 2008, p. 16).

So also on (this interpretation of) the account of Chierchia, Fox and Spector, external exhaustification and the SI of *or* should co-occur.

Summing up, we have seen that all four theories that I presented in the previous chapter make the prediction that SIs and external exhaustification should co-occur. In the next sections, I will present additional results of Experiments 5-7 which tested this prediction. If the results do not confirm the prediction, this casts serious doubt on the exhaustivity view of SIs and therefore on the view that exclusive-*or* is the result of an SI.

3. Experiment 5: Testing co-occurrence of SIs and exhaustivity

3.1 Setup and items

Experiment 5 was already (partly) presented in chapter 4. It was set up in the Possible World Judgment Task (PWJT) paradigm, in which participants were presented with a story containing a dialogue, such as (25). See chapter 4 for a more elaborate description of the paradigm.

(25) Example item Experiment 5
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found at least two marine animals on the beach, that person got to stay up late that night.
The friend said: "Oh, and what had Laura found on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

SIs were assessed by asking participants to answer questions like (26):

(26) Do you think it is possible that Laura had found both a crab and a mussel? yes / no

A 'no' answer to this question indicated that the participant interpreted *or* as exclusive-*or* (*a crab or a mussel but not both*), so with the SI. External exhaustivity was measured by asking participants questions like (27):[6]

(27) Do you think it is possible that Laura had also found something else than a crab or a mussel? yes / no

A 'no' answer to (27) indicated the participant made the external exhaustivity inference that Laura brought nothing else (besides a crab or a mussel).

As I discussed in section 2.1, the crucial prediction was the co-occurrence of the two inferences. I chose not to directly assess this by including both questions in every item, but to include one type of question per item and include both types of items (in all conditions) in every list. This way any effects of the two questions affecting each other within one item

[6] The Dutch phrase for *also something else than A or B* was 'ook nog iets anders dan A of B'.

(due to e.g. answering strategies, carry-over, consistency, order effects) were ruled out. To test the co-occurrence of SIs I compared the SI-rates and the exhaustivity-rates in the same conditions. The rationale of this setup is that if SIs and external exhaustivity come about by one mechanism, they should be triggered to the same extent in a certain environment. Therefore, all conditions should return a similar SI-rate as exhaustivity-rate.

Experiment 5 contained four conditions, in which focus and alternative relevance were manipulated. Example items of the conditions are given in (28)-(31), with the manipulations highlighted in boldface.

(28) Condition 1: [–alternative relevance, +focus]
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **a crab or a mussel** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **what had Laura found** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(29) Condition 2: [+alternative relevance, +focus]
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **at least two marine animals** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **what had Laura found** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(30) Condition 3: [-alternative relevance, -focus]
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **a crab or a mussel** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **who of Laura and Barbara had found a crab or a mussel** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(31) Condition 4: (control)
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found a crab or a mussel on the beach, that person got to stay up late that night.
The friend said: "Oh, and who of Laura and Barbara had found a crab or a mussel on the beach?"
Marieke answered: "**Laura did.**"

Alternative relevance was manipulated by the requirements in the conditionals, and focus was manipulated by the explicit question the target sentence was an answer to (see chapter 4 for details). All four conditions were tested on SIs and exhaustivity (independently). On the view that SIs and external exhaustivity are the result of one mechanism, we expect that in every condition the SI-rate and the external exhaustivity rate will be similar. Notice that the alternative relevance manipulation is the same for the external exhaustivity alternatives (*a crab or a mussel and x*) as for the scalar alternative (*a crab and a mussel*): In condition 1 and 3 (and 4), neither the scalar alternative nor the external exhaustivity alternatives are relevant for the reward, while in condition 2 they both are. Therefore, differences between the SI-rates and exhaustivity rates in a condition cannot be due to differences in the relevance of the alternatives.

3.2 Design, participants and procedure

16 story quadruples like (28)-(31) were created. 12 of these were based on items of Experiment 4. Half of the quadruples were used as SI-items, and the other half as exh-items. The reason for this was that there was a total of 8 conditions (C1-C4 with SI-question and C1-C4 with exh-question), so a design in which the same stories were used in all conditions would require too many participants to yield enough data points per item per condition (for instance 80 participants would only provide 10 data points per item per condition). The 16 quadruples were distributed over 4 lists, so every list contained 8 SI-items (2 per condition) and 8 exh-items (2 per condition). The order of the stories varied between the lists.

68 participants filled out a questionnaire on computers in the lab. (WWSTIM, Veenker 2000). See section 4.3 of chapter 4 for a more elaborate description of the participants and the procedure.

3.3 Results

The SI-rates and exh-rates per condition are presented in Figure 1 and Table 3.

Figure 1: SI- and exh-rates Experiment 5

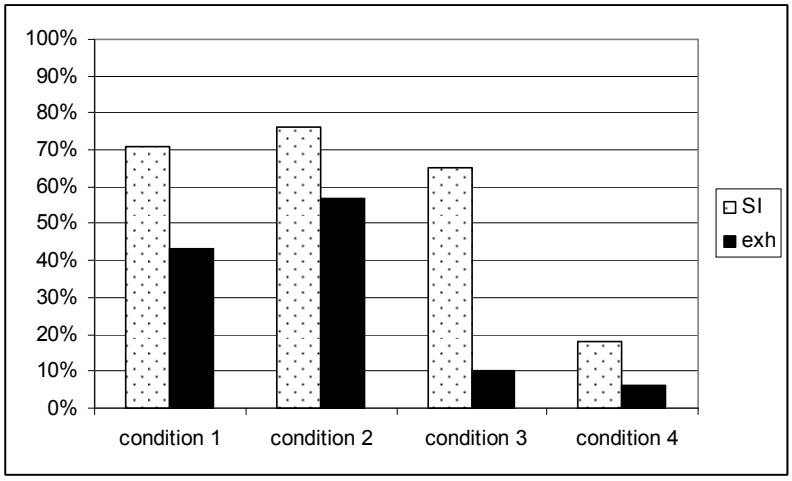


Table 3: SI- and exh-rates Experiment 5

	Condition 1 [-alt +foc]	Condition 2 [+alt +foc]	Condition 3 [-alt -foc]	Condition 4 [control]
SI	71%	76%	65%	18%
exh	43%	57%	10%	6%

Figure 1 clearly show that participants behaved differently on the SI-questions than the exh-questions. A 4x2 ANOVA over participants with SI/exh and condition as (within-subjects) factors revealed a main effect of SI/exh: $F1(1,67) = 76.34$, $p<0.001$, a main effect of condition: $F1(3,201) = 65.88$, $p<0.001$, and an interaction between SI/exh and condition: $F1(3,201) = 16.76$, $p<0.001$. The main effect of SI/exh indicates that over all conditions together, participants behaved differently on the SI-questions than on the exh-questions. The interaction shows that the conditions had a different effect on the exh-questions than the SI-questions, which is obvious from Figure 1. This pattern was confirmed by the items analysis: A mixed ANOVA over items with SI/exh as between-items factor and condition as within-items factor also revealed a significant effect of SI/exh: $F2(1,14) = 86.21$, $p<0.001$ and a significant effect of condition: $F2(3,42) = 84.22$, $p<0.001$. It also returned a significant interaction between condition and SI/exh: $F2(3,42) = 12.22$, $p<0.001$.

Pairwise comparisons showed that in all three test conditions, the SI-questions returned significantly more 'no' answers than the exh-questions: C1: Wilcoxon signed rank test $z = -4.05$, $p<0.001$ (two-tailed), C2: $z = -3.47$, $p = 0.001$ (two-tailed), C3: $z = -6.17$, $p<0.001$ (two-tailed). The control condition (C4) also returned more 'no' answers for the SI-question than for the exh-question: $z = -2.95$, $p = 0.003$ (two-tailed). Apparently, when the target sentence did not contain *or*, participants were more likely to rule out the *both*-situation than the *also something else*-situation based on the stories.

To check whether the differences between the SI-rates and the exh-rates in the test conditions were bigger than this baseline-difference, for each test condition a 2x2 ANOVA pair (with test condition/control condition as factor and SI/exh as factor) was calculated with the SI and exh data of this test condition and those of the control condition 4. As in the omnibus ANOVAs, in the participants ANOVAs test/control and SI/exh were within-subjects factors, and in the items ANOVAs SI/exh was a between-items factor and test/control was a within-items factor. A significant interaction of test/control and SI/exh was observed for condition 1: $F1(1,67) = 5.75$, $p = 0.019$, $F2(1,14) = 6.67$, $p = 0.022$. There was no significant interaction for condition 2: $F1(1,67) = 1.67$, $p = 0.201$, $F2(1,14) < 1$. Finally, condition 3 did reveal a significant interaction of test/control and SI/exh: $F1(1,67) = 39.41$, $p<0.001$, $F2(1,14) = 20.55$, $p<0.001$. So in condition 1 and condition 3, the difference between the SI-rate and the exh-rate was bigger than in the baseline condition without *or* (C4). This goes against the prediction that the SI-rates and the external exhaustivity should be similar.

3.4 Discussion

The data show that in the same environments (the experimental conditions) SI-rates and exh-rates were not similar: more SIs were calculated than external exhaustifications in all conditions. Even if we take the difference in the baseline-condition 4 into account, the interactions revealed that in two of the three test conditions, still more SIs were triggered than exhaustifications. This goes against the view that SIs and external exhaustivity are the result of one mechanism. Furthermore, the data clearly show that the experimental manipulations had a different effect on SIs than exhaustivity, which is further support for the view that they are not the result of one mechanism.

However, we have to be careful in drawing conclusions from these data. As I pointed out in section 7.2 of chapter 4, the effect of the manipulations of focus and alternative relevance on the SI-questions might have been masked as the experimental question might have overruled the manipulations. I

discuss how this relates to the current data in the general discussion in section 5 below.

There was one factor in the experiment which was not controlled for. As different stories were used to test for SIs and exhaustivity, it cannot be ruled out that the difference between SI-rates and exh-rates is due to other differences between the stories. Perhaps SIs and exhaustivity-inferences do always co-occur, but the stories in the exhaustivity conditions were such that they were less likely to trigger the SI/exhaustivity inferences. One argument against this is that 12 out of the used 18 story quadruples (all 8 SI-story quadruples and half (4) of the exhaustivity story quadruples) were based on items of Experiment 4, discussed in chapters 4 and 8. So half of the exhaustivity story quadruples were based on items of Experiment 4, where they triggered similar SI-rates as observed for the other 8 which were re-used as SI-items in Experiment 5. But perhaps the other half of the exhaustivity items, which were new items, were deviant. A quick look at the data shows that this is not the case. Both the re-used items and the new items trigger the same behavior. See Table 4 for the exhaustivity percentages of the re-used stories versus the new stories and the total.

Table 4: Exhaustivity data split in re-used stories and new stories

	Condition 1	Condition 2	Condition 3	Condition 4
re-used	46%	60%	15%	6%
new	40%	54%	6%	6%
total	43%	57%	10%	6%

However, to make sure the differences between SIs and external exhaustivity were not due to differences between the two sets of stories, I conducted a follow-up experiment in which one set of stories was used to test both SIs and external exhaustivity, in a between-subjects design. This was the first of two paper-and-pencil control experiments: Experiments 6 and 7, which I will discuss in the next section.

4. Experiments 6 and 7: Paper-and-pencil control experiments

4.1 Experiment 6: Using the same stories for SI and exh items

4.1.1 Setup, items, design, participants and procedure

In Experiment 6 the items and design of Experiment 4 were used, which were very similar to those of Experiment 5 (see 5.1 of chapter 4 for the minor differences between the two sets). Two versions of the questionnaire were created. One version was identical to Experiment 4 and contained only SI-questions, such as (26) above. In the other version the SI-questions were replaced by exh-questions like (27) above. Four lists were created of both versions, in the same way as in Experiment 5.

Participants were undergraduate students of Utrecht University. They filled out a paper-and-pencil questionnaire in class. All participants were native speakers of Dutch and had no prior knowledge of the topic. They were students of the Faculty of Humanities and all were taking the same (unrelated) course. One group of students filled out the SI-version of the experiment, while another group filled out the exh-version. Due to differences in the sizes of the classes, 63 students filled out the SI-version, while 35 filled out the exh-version. The instructions were the same as in Experiment 5, with the additional instruction that participants were discouraged to look back at earlier items. This instruction was included in order to increase similarity to the computer-based experiments, in which items were presented one by one on the screen, with no option of going back to earlier items.

4.1.2 Results

The results were very similar to Experiment 5, see Figure 2 and Table 5:

Figure 2: SI- and exh-rates Experiment 6

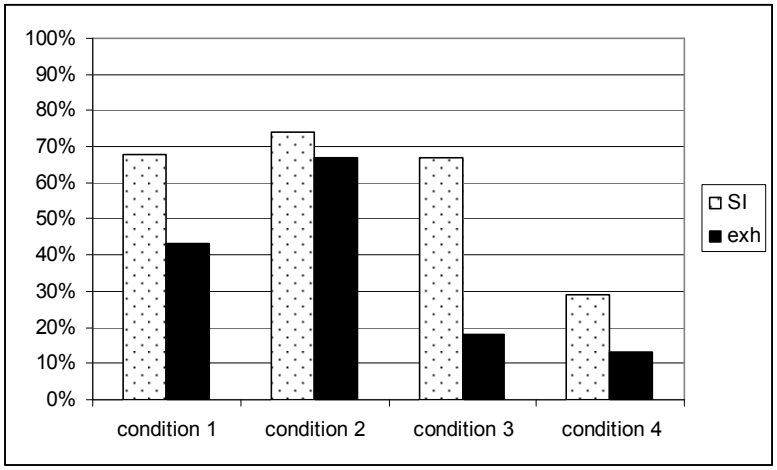


Table 5: SI- and exh-rates Experiment 6

	Condition 1 [-alt +foc]	Condition 2 [+alt +foc]	Condition 3 [-alt -foc]	Condition 4 [control]
SI	68%	74%	67%	29%
exh	43%	67%	18%	13%

A mixed ANOVA over participants with SI/exh as between-subjects factor and condition as within-subjects factor revealed a significant effect of SI/exh: $F1(1,96) = 22.62$, $p<0.001$, a significant effect of condition: $F1(3,288) = 51.39$, $p<0.001$ and a significant interaction of SI/exh and condition: $F1(3,288) = 9.10$, $p<0.001$. A 4x2 ANOVA over items with SI/exh and condition as within-items factors returned the same result: main effect of SI/exh: $F2(1,11) = 65.01$, $p<0.001$, main effect of condition: $F2(3,33) = 34.80$, $p<0.001$, interaction of SI/exh and condition: $F2(3,33) = 7.89$, $p<0.001$.

Pairwise comparisons showed significant differences between the SI-version and the exh-version in condition 1 (Mann-Whitney $U = 732.00$, $z = -2.87$, $p = 0.004$ (two-tailed)), condition 3 ($U = 381.00$, $z = -5.56$, $p<0.001$ (two-tailed)), and condition 4 ($U = 810.50$, $z = -2.37$, $p = 0.018$ (two-tailed)), but not in condition 2 ($U = 948.00$, $z = -1.24$, $p = 0.216$ (two-tailed)).

As the SI-rate was again higher than the exh-rate in the baseline condition, we again have to consider the 2x2 ANOVA pairs to see whether the difference was bigger for the test conditions than the baseline condition. In the subjects analysis, SI/exh was treated as a between-subjects factor and condition as a within-subjects factor and in the items analysis, both SI/exh

and condition were treated as within-items factors. The ANOVAs did not reveal a significant interaction of SI/exh and control/test for condition 1 (with condition 4): $F1(1,96)<1$, $F2(1,11)<1$, but there was a significant interaction of SI/exh and test/control for condition 3: $F1(1,96) = 16.44$, $p < 0.001$, $F2(1,11) = 20.96$, $p = 0.001$. We can conclude that in condition 3 the difference between the SI-rate and the exh-rate was bigger than in the baseline condition. So again the prediction of co-occurrence of SIs and exhaustivity was contradicted by the data.

4.1.3 Discussion

Experiment 6 replicated the results of Experiment 5 in a design in which the same stories were used to trigger SIs as external exhaustivity. In conditions 1 and 3 more SIs were calculated than exhaustifications and the difference in condition 3 was significantly bigger than the baseline difference. This rules out the explanation that the differences between SI-rates and exh-rates in Experiment 5 were caused by differences between the stories.

As Experiment 6 was a paper-and-pencil experiment with the stories and design of Experiment 4, I decided to also do paper-and-pencil version with the items and design of Experiment 5. The goal of this experiment was just to complete the picture.

4.2 Experiment 7: Paper-and-pencil version of Experiment 5

4.2.1 Setup, items, design, participants and procedure

Experiment 7 was a paper-and-pencil version of Experiment 5, so again both SIs and exhaustivity were tested, but like Experiment 5 in a within-subjects design (with different stories for SI and exh). While Experiment 6 used the items of Experiment 4, here the items of Experiment 5 were used. The design was also adopted from Experiment 5. 34 undergraduate students of Utrecht University filled out the questionnaire after an exam. None of them had participated in any of the previous experiments or had any prior knowledge of the topic, and all of them were native speakers of Dutch.

4.2.2 Results

The results of Experiment 7 are given in Figure 3 and Table 6.

Figure 3: SI- and exh-rates Experiment 7

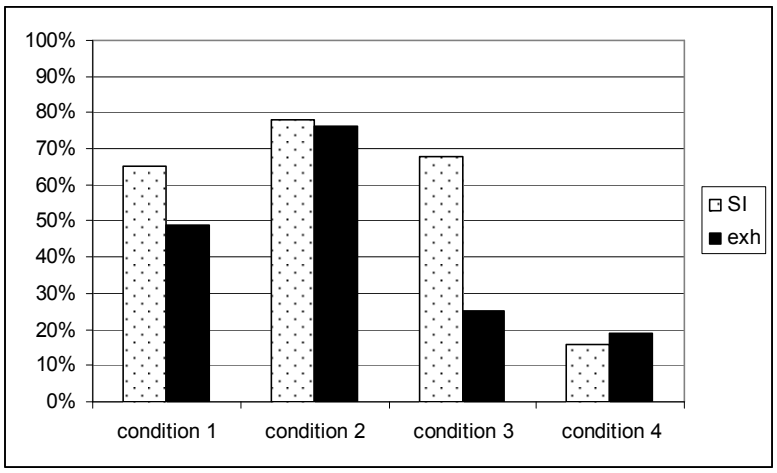


Table 6: SI- and exh-rates Experiment 7

	Condition 1 [-alt +foc]	Condition 2 [+alt +foc]	Condition 3 [-alt -foc]	Condition 4 [control]
SI	65%	78%	68%	16%
exh	49%	76%	25%	19%

A 4x2 ANOVA over participants and a mixed ANOVA over items with SI/exh and condition as factors showed a significant effect of SI/exh: $F1(1,33) = 11.69$, $p = 0.002$, $F2(1,14) = 36.94$, $p<0.001$, a significant effect of condition: $F1(3,99) = 35.58$, $p<0.001$, $F2\ (3,42) = 29.81$, $p<0.001$ and a significant interaction of SI/exh and condition: $F1(3,99) = 8.50$, $p<0.001$, $F2\ (3,42) = 4.90$, $p = 0.005$.

Pairwise comparisons returned no significant differences between the SI-rates and exh-rates of condition 1 (Wilcoxon signed rank test $z = -1.80$, $p = 0.072$ (two-tailed)), condition 2 ($z = -0.166$, $p = 0.868$ (two-tailed)), and condition 4 ($z = -0.63$, $p = 0.527$ (two-tailed)). However, a significant difference was observed in condition 3: $z = -3.77$, $p<0.001$ (two-tailed). As in the baseline condition the SI-rate is not higher than the exh-rate, we know the difference in condition 3 is due to a difference in SI and exhaustivity inferences. So again the co-occurrence of SIs and exhaustivity was contradicted.

5. General discussion and conclusions

The results of Experiments 5-7 are summarized in Table 7:

Table 7: Summary SI-rates and external exhaustivity-rates Exp. 5-7

	Condition 1 [−alt +foc]		Condition 2 [+alt +foc]		Condition 3 [−alt −foc]		Condition 4 [control]	
	SI	exh	SI	exh	SI	exh	SI	exh
Exp 5	71%	43%	76%	57%	65%	10%	18%	6%
Exp 6	68%	43%	74%	67%	67%	18%	29%	13%
Exp 7	65%	49%	78%	76%	68%	25%	16%	19%

In two out of three test conditions in Experiment 5 (C1 and C3) and one of three in Experiments 6 and 7 (C3), significantly more SIs were calculated than external exhaustivity-inferences, even after we controlled for the baseline difference. Individual differences between participants or stories cannot have been responsible for the difference, as the same participants answered SI- and exh-questions in Experiments 5 and 7 and the same stories were used for SI- and exh-items in Experiment 6. The results of Experiments 5-7 suggest that SIs and external exhaustification do not come about by the same mechanism, which goes against the theories discussed above.

However, in section 7.2 of chapter 4 I pointed out that the reason that the effect of focus (and alternative relevance) on SIs was absent or small in Experiments 4-7 might be because the experimental question overruled the critical manipulations. I argued that asking explicitly about the *A and B* situation provided alternative relevance and focus in all three conditions. Furthermore, as the experimental question was the most recent issue that was raised when participants (re)interpreted the target sentence, it became the question relative to which the target sentence was interpreted. This resulted in similar SI-rates in the three test conditions. This explanation was supported by the reading times data of Experiment 9. There, I used similar stories to those of condition 2 ([+alternative relevance, +focus]) and condition 3 ([−alternative relevance, −focus]) of Experiments 5-7, and clear indications were found (particularly on the plural pronoun *them*) that more SIs were calculated in the [+alternative relevance, +focus] condition.

However, the data presented in this chapter suggest the masking was absent for the exhaustivity data, which showed great sensitivity to the manipulations. In order to explain this difference in masking I will consider a number of differences between the SI-items and the exh-items which might account for this. However, these differences alone might explain the observed difference between SI-rates and exh-rates, making it impossible to draw conclusions about whether SIs and exhaustivity are the result of one

mechanism. Therefore, I will leave the discussion of the differences open. This way the reader can decide whether the observed difference between the SI-rates and the exh-rates is meaningful (and the masking hypothesis in chapter 4 should be rejected) or whether the difference between SI-rates and exhaustivity rates is completely due to a difference in masking of the effect by the experimental paradigm, or whether the truth is in the middle.

A possible difference between the SI-conditions and the exh-conditions is that the relation between the experimental question and the target sentence was different. I repeat the two types of experimental questions here for convenience.

(26) SI-question
Do you think it is possible that Laura had found both a crab and a mussel? yes / no

(27) exh-question
Do you think it is possible that Laura had also found something else than a crab or a mussel? yes / no

The target sentence in both cases was *Laura found a crab or a mussel*. If we only take the experimental question and the target sentence into account, there is an obvious difference between the two. As *A or B* was used in the target sentence, relative to the SI-question in (26) there was overt material in the target sentence which contrasted with the questioned situation (*A and B*). This might have caused participants to look no further and judge the situation impossible. Relative to the exh-question in (27), there was no overt material in the target sentence that contrasted with the questioned situation. The exh-question was whether something else than A or B was also found, while the sentence itself contained just *A or B*. Therefore, participants had to turn to the wider context for clues. As the wider context contained the critical manipulations, it is not surprising that the manipulations had a stronger effect in the exhaustivity conditions. In the most extreme case, in the SI-conditions participants might have resorted to an answering strategy along the lines of: 'When *or* is in the answer, the *both*-situation is not possible.' This strategy was useless in the external exhaustivity conditions, and participants turned to the wider context for clues. Something that might have added to this difference was the wording of the experimental question. In the external exhaustivity question in (27) I used the phrase *also something else than* (Dutch: 'ook nog iets anders dan'). Participants might have taken this as a cue to go beyond what the speaker said and look at the wider context, while in the SI-question this cue was absent. Therefore, an

interesting alternative exhaustivity-question would have been asking directly about another alternative, say *B and C*, see (27'):

(27') Do you think it is possible that Laura had found a mussel and an oyster? yes / no

However, my guess is that this question would still trigger less 'no' answers (at least in condition 3) than the SI-question in (26), or the version of it that is minimally different from (27'):

(26') Do you think it is possible that Laura had found a crab and a mussel? yes / no

Anyway, as I did not test these questions, this remains speculation (and a suggestion for further research).

Summing up, the fact that the experimental question was different between the SI-items and the exh-items could explain why the contextual manipulations were more successful for the exh-items than for the SI-items. This explains the difference in the size of the masking effect of the experimental question overruling the critical manipulations: Due to the difference in the relation between the experimental question and the target sentence, the overruling effect was stronger in the SI-conditions.

We could conclude the exhaustivity account of SIs is not contradicted and the difference observed is just experimental noise. However, let us make the assumption that SIs and external exhaustivity are indeed the result of one mechanism and look at the above explanation. I first proposed that the use of *or* created a contrast with the questioned *A and B* situation. However, this contrast is only present on the exclusive reading of *or*. On the inclusive reading of *or*, the questioned situation is possible. On the view that the exclusive reading comes about by application of an exhaustivity operator, we have to assume the contrast came about by application of this operator, or by participants recognizing that applying the operator would lead to a contrast. However, then the question arises why this did not happen in the exhaustivity conditions. There, application of the exhaustivity operator would also create a contrast between the sentence and the questioned situation. The sentence would then mean *A or B and nothing else*, while the questioned situation is *A or B and X*. The point is that although it looks obvious that the two questions gave rise to different behavior, it is not so obvious anymore once we take the view that exclusive-*or* comes about by an exhaustivity operator.

As I said, I will not be able to draw strong conclusions. On the one hand, the difference between the SI-rates and the exh-rates is striking and

goes against the views presented above. On the other hand, the differences between the experimental questions that were used to assess the two types of inferences might have been responsible for the difference. However, as I pointed out above, these differences are not so obvious from the point of view of the exhaustivity theories. I conclude the data at least raise doubts about whether the SI of *or* is the result of application of an exhaustivity operator, although additional experimental support is needed to support this claim. Irrespective of this, it seems in general exhaustivity is much more sensitive to context than exclusive-*or*. Therefore, next to the problem of the Speaker Expertise Paradox in the previous chapter, this is another reason to doubt the view that exclusive-*or* comes about by an SI. I will return to this in the next chapter, in which I summarize the thesis and draw conclusions. In section x I consider the conclusions for theories about how SIs are derived. I argue none of the current theories can account for the complete range of data presented in this thesis, and I briefly consider some alternative views of *or*.

CHAPTER 10: SUMMARY AND CONCLUSIONS

1. Introduction

In this dissertation I set out to contribute to a more explicit theory of parts of utterance meaning that are traditionally assumed to belong to the realm of linguistic pragmatics. Through the wastebasket function of pragmatics, these parts of utterance meaning have been accounted for far less systematically than parts of meaning that are considered to belong to truth-conditional compositional semantics. Many theories in pragmatics propose very general principles and illustrate these by examples that trigger intuitions which can be explained by the general principles. However, theories that make clear predictions about the factors at work are quite rare. The context-sensitivity of Scalar Implicatures (SIs) is a case in point. SIs, for instance the inference that from *John has A or B* follows that John does not have A *and* B, based on the entailment scale <or, and>, are considered to be one of the strongest and most predictable types of pragmatic inference. However, it is widely acknowledged that whether or not they arise depends on the specific context the scalar term appears in. While much is known about the properties at the sentence level that trigger or block SIs, little is known about the properties of the wider context that lead to their presence or absence. The contextual effects are usually described in very general terms or they are set aside as special circumstances that are of secondary interest. I argued that the question in which contexts one of the most robust pragmatic inferences arises, is a fundamental question for an explicit theory of pragmatics.

In the search for contextual properties that affect SI-calculation I took an experimental approach. In determining the complete meaning of an utterance, including situational and contextual inferences, many factors are at work at the same time. Therefore, we cannot provide conclusive evidence for a certain property through examples that trigger certain intuitions. However, through well-controlled experiments we can isolate a certain property and reveal its effects. The experimental approach also allows us to assess the psychological reality of theoretical constructs, as in the end we (or at least most linguists) want to describe the language system of the actual language user. Furthermore, empirical research adds to the ever-growing pool of data that survives when theories die, and that future theories can build on.

However, there are exceptions to the rule of describing context-sensitivity of SIs without making testable predictions or ignoring it altogether. I tested the predictions of two of these exceptions, the accounts of Van Kuppevelt (1996) and Van Rooij (2002). Both make the explicit

prediction that SIs should only arise if the scalar term is part of the information focus of the sentence, i.e. the part that contains new information. Van Kuppevelt's claim is based on a framework in which discourse is analyzed as an ongoing questioning process. Van Rooij proposes an explicit account of how SIs are calculated. According to him they are calculated through the application of an exhaustivity operator. This operator selects the minimal meaning of the information focus of the sentence, which is how it gives rise to the SIs. As the information focus is defined as the new information, which part of the sentence is the information focus is determined by the prior context. Therefore, on the view that SIs only arise in the focus, the prior context determines whether an SI arises in a particular sentence. Hence, information focus is a contextual property that determines when SIs arise.

A clear example of how context determines focus, and therefore SIs, is through questions. These questions can be explicitly given or can be implicitly triggered by the context. In Experiments 1-9 (with the exception of Experiment 3), I tested explicit question-answer pairs like (1) and (2):

(1) Q: What did Harry bring?
A: Harry brought bread or chips$_F$.

(2) Q: Who brought bread or chips?
A: Harry$_F$ brought bread or chips.

In (1) the SI that Harry did not bring bread and chips is predicted to arise, as the scalar term *or* is in the part of the sentence that contains the new information, the focus. In (2) the SI is predicted not to arise, as *or* is part of the background.

Chapters 3-7 tested this prediction in different paradigms. Remember that we can only determine the crucial properties at work in pragmatic inference by isolating them in well-controlled experiments. Therefore, the methodological goal was to find a suitable paradigm to test the effect of focus on SIs, and to assess SIs in general. I started out with the most widely used paradigm to test SIs, the Truth Value Judgment Task (TVJT). Three experiments were carried out in this paradigm, which I summarize in section 2. The main disadvantage of this paradigm is that it requires the actual situation to be given. This is contrary to how SIs arise in everyday conversation, as inferences about what actually happened. Therefore, I tested the hypothesis in a new experimental paradigm, the Possible World Judgment Task (PWJT). In this paradigm the actual situation is left out, creating a more natural environment for SIs. I summarize these experiments in section 3. However, in the PWJT the experimental question overruled the

contextual manipulation of focus. Therefore, I switched to on-line processing experiments, in which SIs could be assessed implicitly, through the processing cost they bring along. In these experiments I adopted the Self-Paced Reading (SPR) paradigm for measuring reading times. I present these experiments in section 4. All experiments up to that point tested question-answer pairs like (1) and (2) above, where the focus was determined by wh-questions and the scalar term *or* appeared in direct object position. In the experiments presented in chapter 7 I broadened the scope of the investigation to another scalar term (*most*) in another scalar position (subject) and I considered answers to another type of question (yes/no-questions). These experiments will be summarized in section 5.

In chapters 8 and 9 I investigated the psychological reality of the claim that the exclusive reading of *or* comes about by an SI. It is well known in the literature that the Gricean derivation of SIs is problematic for *or*. *Or* typically triggers the inference that the speaker is not completely informed, which goes against one of the crucial steps in the Gricean derivation of SIs, the assumption of speaker expertise. In chapter 8 I presented additional data from a PWJT experiment that address this issue. In section 6 I summarize these findings, as well as the solutions that have been proposed in the literature. In chapter 9 I investigated the relation between SIs and exhaustivity. One of the theories on which the prediction of focus-sensitivity of SIs was based was Van Rooij's (2002) view that SIs are the result of an exhaustification procedure. This theory predicts that for *A or B*, the SI that *A and B* is not the case should always co-occur with the external exhaustivity inference that *C* is not the case. I argued this prediction does not just follow from the view of Van Rooij, but also from the grammatical exhaustivity accounts of Chierchia and colleagues. I presented additional data from the PWJT experiments which tested this claim. These data are summarized in section 7.

In the summaries of the experiments in sections 2-7 I will provide example items of all of them to provide a clear overview of the development through the experiments. All example items are translated from their original Dutch counterparts. Finally, in sections 8-10 I draw conclusions and provide directions for future research on the three topics that were central in this thesis: the focus-sensitivity of SIs, the methodology for assessing SIs and the psychological reality of the SI-view of *or*.

2. The TVJT Experiments (chapter 3)

In chapter 3 the prediction of focus-sensitivity of SIs was tested with a number of Truth Value Judgment Task (TVJT) experiments. Contrary to the

normal setup of a TVJT experiment, the target sentence was not presented in isolation but as an answer in a dialogue as in (1) and (2) above. The story described the *A and B* situation, where in the target sentence *A or B* was used. Therefore, if participants judged the answer containing *A or B* 'false', this indicated they calculated the SI.

In Experiment 1 I manipulated the focus structure of the target sentence by varying the explicit question between conditions. Example items are given in (3) and (4), translated from Dutch:

(3) Experiment 1 focus condition
Katja was searching for marine animals on the beach near her grandparents' house. She had promised her grandfather to find some beautiful animals. He had said that if she would find an oyster, she would get ten bucks. Katja soon found a crab. Not much later she also found a starfish. But no matter how hard she looked, she didn't find an oyster.

A: "What did Katja find?"
B: "Katja found a crab or a starfish."

(4) Experiment 1 non-focus condition
Katja and Birgit were searching for marine animals on the beach near their grandparents' house. Their grandfather had encouraged them both to go look for a crab or a starfish. He had promised them that the one who would find a crab or a starfish, would get ten bucks. After some searching Katja found a crab. Not much later she also found a starfish. Birgit couldn't find anything and had to return to the house empty handed.

A: "Who found a crab or a starfish?"
B: "Katja found a crab or a starfish."

I also varied the set sizes of the subject set (one person in (3) versus two persons in (4)) and direct object sets (three objects in (3) versus two objects in (4)), in order to support the question that was asked by speaker A. The material in the second and third sentences was also varied between conditions to provide support for the question. The results were that in the focus condition the answer was considered false in 73% of the cases, versus 55% in the non-focus condition, which was a significant difference. There probably was some carry-over from the focus condition to the non-focus condition, because if we consider only the very first test items participants were presented with, the focus condition was considered false in 71% of the

cases, but the non-focus condition only in 37% of the cases. These results indicate that SIs are indeed sensitive to focus.

In Experiment 2 I reduced the differences between the two conditions to the bare minimum: the focus-determining explicit question. This way we could be sure the effect found was due to the question, and not to other differences between the conditions. I also presented the question-answer pair in indirect speech, as in Experiment 1 the relation between the speakers and the story was unclear. Items were for instance (5) and (6). I mark the differences between the two conditions in boldface for easy reference (not in the actual items).

(5) Experiment 2 focus condition
Julie and Karin were searching for marine animals on the beach. After some searching Julie found a crab. Not much later she also found a starfish. Unfortunately, Karin didn't find anything.
When **Karin** returned, her mother asked **what kind of marine animals Julie had found**.
Karin answered that Julie had found a crab or a starfish.

Is Karin's answer true? true / false

(6) Experiment 2 non-focus condition
Julie and Karin were searching for marine animals on the beach. After some searching Julie found a crab. Not much later she also found a starfish. Unfortunately, Karin didn't find anything.
When **they** returned, their mother asked **who had found a crab or a starfish**.
Karin answered that Julie had found a crab or a starfish.

Is Karin's answer true? true / false

Again, significantly more 'false' answers were given in the focus condition: 67%, versus 41% in the non-focus condition. The difference was again bigger for the first items: 61% versus 26%. As only the focus-determining question was varied between conditions, these results clearly show that more SIs are calculated for *or* if it is part of the information focus of the sentence than if it is part of the background.

In order to show that this effect was not due to the repetition of *or* from the question, I took out the explicit question in a follow-up experiment (Experiment 3). To make sure participants assigned the right focus structure to the target sentence, I used spoken stimuli in which the stress pattern of the sentences was manipulated, based on the rule that the main stress of a

sentence falls on (a syllable of) the information focus. I also increased the set size difference used in Experiment 1 to contextually trigger the implicit question that matched the focus structure of the target sentence. Consider example items (7) and (8), where capitals indicate the main stress of the sentence.

(7) Experiment 3 focus condition
Paola was getting lunch at the cafeteria. At the fruit section there were bananas, oranges, apples and pears.
Bananas Paola didn't like, so she didn't take any.
The oranges looked a bit old, so she also didn't take any of those.
The apples looked nice and juicy, so Paola took one.
A pear Paola hadn't eaten in years, so for a change she decided to also take a pear.

"Paola took AN APPLE OR A PEAR from the fruit section."

(8) Experiment 3 non-focus condition
Paola, Linda, Betty and Ginger were getting lunch at the cafeteria. At the fruit section there were only apples and pears.
Paola was a real health freak, so she took an apple. After some consideration she also took a pear.
Linda already got lots of other food, so she decided not to take any fruit.
Betty didn't like fruit at all, so she ignored the apples and the pears.
Ginger considered taking a pear, but she wasn't sure whether she brought enough money, so she didn't take any.

"PAOLA took an apple or a pear from the fruit section."

The observed SI-rates were very similar to those of Experiments 1 and 2: 85% for the focus condition and 55% for the non-focus condition. I concluded the focus-sensitivity of SIs was also confirmed when there was no explicit question, and therefore it was not an effect of the repetition of *or* from the question. The results of Experiments 1-3 are summarized in Table 1, with the data of only the first items added between parentheses.

Table 1: Summary SI-rates Experiments 1-3

	focus	nonfocus
Exp 1	73% (71%)	55% (37%)
Exp 2	67% (61%)	41% (26%)
Exp 3	85%	55%

Taking together the results of Experiments 1-3, we see that there was a robust effect of focus on calculation of the SI of *or*. However, the observed percentages were not as black-and-white as predicted by the focus-sensitive theories. Three possible explanations were given in chapter 3: chance performance in the non-focus condition, a mismatch between implicit and explicit question in the context, and the default position of information focus at the end of the sentence. But perhaps the strongest argument against attributing too much importance to the gradual instead of categorical nature of the contrast is that the TVJT due to its very nature is not very suitable to assess SIs. As the actual situation is already known when the scalar term has to be interpreted, this is a very unnatural environment for SI-calculation. For this reason I used a new paradigm in chapter 4, the PWJT.

3. The PWJT Experiments (chapter 4)

In the PWJT experiments I considered an additional contextual property that has been claimed to affect SI-calculation. One of the key steps in the nonce derivation by which SIs are calculated according to the standard Gricean view (see chapter 1), is the assumption of the hearer that the stronger scalar alternative is relevant ('alternative relevance' from now on). For a sentence like *Katja found a crab or a starfish*, the stronger scalar alternative is the same sentence with *or* replaced by *and*: *Katja found a crab and a starfish*. As in the items of the TVJT-experiments it was unclear why the question in the story was asked, the participants might have guessed what was of interest to the hearer (which in those cases was the questioner), based on the question she asked. As it is more likely that the stronger scalar alternative was of interest to her if she asked *What did Katja find?* (the question in the focus condition) than if she asked *Who found a crab or a starfish?* (the question in the non-focus condition), this could explain the effect on SI-calculation in Experiments 1-3 that we attributed to focus.

In Experiments 4-7, I tried to tease apart the effects of focus and alternative relevance, by including three conditions: one with both focus and alternative relevance, one with neither and crucially, one with focus but without alternative relevance (condition 1 below). Alternative relevance was manipulated through conditionals in the context in which a requirement was given, for which the stronger scalar alternative either was relevant or not. For instance, for the requirement *if Katja found a crab or a starfish* the stronger scalar alternative is not relevant, as finding one of the two is enough, but for the requirement *if Katja found at least two animals*, the stronger scalar alternative is relevant.

In order to create these conditions and avoid the problem of the TVJT that the actual situation was given, I designed a new story type. By introducing a story-teller and making the questioner the person to whom the story was told, and by taking out the actual situation of who found what from the story, it was clear the questioner asked the question in order to find out whether the requirement of the conditional was satisfied. Consider the example items of Experiment 4 in (9)-(11), with differences again highlighted in boldface:

(9) Experiment 4 condition 1: [−alternative relevance, +focus]
Marieke told her mother that Laura went searching for marine animals on the beach yesterday, and that her father had told her that **if she would find a crab or a mussel**, she would get to stay up late that night.
The mother said: "Oh, and **what did Laura find?**"
Marieke answered: "She found a crab or a mussel."

(10) Experiment 4 condition 2: [+alternative relevance, +focus]
Marieke told her mother that Laura went searching for marine animals on the beach yesterday, and that her father had told her that **if she would find at least two marine animals**, she would get to stay up late that night.
The mother said: "Oh, and **what did Laura find?**"
Marieke answered: "She found a crab or a mussel."

(11) Experiment 4 condition 3: [−alternative relevance, −focus]
Marieke told her mother that Laura **and Barbara** went searching for marine animals on the beach yesterday, and that their father had told them that **the one who would find a crab or a mussel**, would get to stay up late that night.
The mother said: "Oh, and **who of them found a crab or a mussel?**"
Marieke answered: "Laura found a crab or a mussel."

In order to assess the SIs now that the actual situation was taken out, I used a new paradigm which I called the *Possible World Judgment Task (PWJT)*, in which participants were asked to judge the possibility of the *A and B* situation, as in (12):

(12) Do you think it is possible that Laura found both a crab and a mussel? yes / no

A 'no' answer to this question meant the participant considered the *A and B* situation to be impossible, which indicated she calculated the SI. The advantage of this paradigm was that as the actual situation was not given, it provided a much more natural environment to calculate the SI.

With the actual situation taken out, I had to control for plausibility effects. Therefore, I included a control condition which was identical to condition 3, but in which the term answer was given and the VP containing *or* was elided, see (13). If 'no' answers were due to implausibility, this condition should also return a high rate of 'no'-answers.

(13) Experiment 4 condition 4: (control)
Marieke told her mother that Laura and Barbara went searching for marine animals on the beach yesterday, and that their father had told them that the one who would find a crab or a mussel, would get to stay up late that night.
The mother said: "Oh, and who of them found a crab or a mussel?"
Marieke answered: "**Laura did**."

Based on the TVJT-results and the claims about importance of alternative relevance, we expected more SIs in the [+alternative relevance, +focus] condition 2 than the [–alternative relevance, –focus] condition 3. The crucial condition for teasing apart focus and alternative relevance was condition 1: if it returned a high SI-rate like condition 2, focus was more important than alternative relevance. However, if its SI-rate was low like condition 3, this meant alternative relevance was more important. A third possibility was that the SI-rate of condition 1 was in between the rates of condition 2 and condition 3, which would suggest both properties are important for SIs.

The SI-rates of Experiment 4 were as follows: Condition 1: 67%, condition 2: 72%, condition 3: 63%, condition 4: 27%. However, none of the differences between the three test conditions (conditions 1-3) were significant, which made it difficult to draw conclusions. Therefore, I conducted a follow-up experiment (Experiment 5) in which three possible sources of experimental noise in Experiment 4 were avoided: The first was the fact that there were more differences between the conditions than strictly needed for the critical manipulation (e.g. different set sizes, different forms of the conditional). The second source of experimental noise was the speaker expertise question which preceded the SI-question (which was discussed in chapter 8, see section 6 below). The third issue was that the requirement of e.g. finding at least one of the two animals in conditions 1 and 3 might have made it implausible that the characters in the story went on searching and found the other one (although this was not reflected in the control condition).

These problems were all avoided in Experiment 5, of which example items are given in (14)-(17):

(14) Experiment 5 condition 1: [–alternative relevance, +focus]
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **a crab or a mussel** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **what had Laura found** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(15) Experiment 5 condition 2: [+alternative relevance, +focus]
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **at least two marine animals** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **what had Laura found** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(16) Experiment 5 condition 3: [–alternative relevance, –focus]
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **a crab or a mussel** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **who of Laura and Barbara had found a crab or a mussel** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(17) Experiment 5 condition 4: (control)
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found a crab or a mussel on the beach, that person got to stay up late that night.
The friend said: "Oh, and who of Laura and Barbara had found a crab or a mussel on the beach?"
Marieke answered: "**Laura did.**"

The results were similar to those of Experiment 4: condition 1 yielded 71% SIs, condition 2 76%, condition 3 65% and condition 4 18%. Here, the difference between condition 2 and condition 3 was significant, indicating that focus and alternative relevance together increased SIs. However, condition 1 did not differ significantly from either condition 2 or condition 3, so it was impossible to determine the relative importance of focus and alternative relevance for SI-calculation.

Experiments 6 and 7 were paper-and-pencil versions of Experiments 4 and 5, and returned similar SI-rates. Again, the differences were small. Nevertheless, in Experiment 6 the difference between condition 2 and condition 3 was significant, but only on the subjects analysis. In Experiment 7 the difference between condition 1 and condition 2 was significant (but again, only over subjects), suggesting an effect of alternative relevance, but other differences were not significant. Overall, the differences between the conditions were small. The results of Experiments 4-7 are summarized in Table 2.

Table 2: Summary SI-rates Experiments 4-7

	Condition 1 [-alt, +foc]	Condition 2 [+alt, +foc]	Condition 3 [-alt, -foc]	Condition 4 [control]
Exp 4	67%	72%	63%	27%
Exp 5	71%	76%	65%	18%
Exp 6	68%	74%	67%	29%
Exp 7	65%	78%	68%	16%

All we can conclude from Table 2 is that the results suggest that focus and alternative relevance together increase SIs, based on the significant differences between conditions 2 and 3 in Experiments 5 and 6. However, the differences between condition 1 and the other test conditions are too small to draw conclusions about their mutual relation.

I argued that the reason that the differences were much smaller than in the TVJT might have been that the experimental question (e.g. *Do you think it is possible she found both?*) overruled the critical manipulations. Results from the literature show that asking a participant explicitly for a certain interpretation, makes this issue relevant (see section 7.2 of chapter 4). As the experimental question always targeted the situation which could have been described by the stronger scalar alternative, this provided alternative relevance in all conditions. Similarly, the question always implied the direct object, (e.g. what Laura found) and relative to this question *A or B* was always in the information focus. So the experimental question overruled the critical manipulations, turning all conditions into [+alternative relevance, +focus] conditions. This problem did not arise in the TVJT, which asked for

a truth value instead of a particular interpretation. Also, in the PWJT the experimental question was the most recent issue when participants (re)interpreted the target sentence. This is contrary to the TVJT, where the focus-determining question (which contained the critical manipulation) was the most recent issue. In that light it is not surprising the differences between the conditions were smaller in the PWJT-experiments.

Therefore, I decided we needed a paradigm which has best of both worlds (of the TVJT and the PWJT): a paradigm in which the actual situation as well as the experimental question about the *A and B* situation could be left out. Therefore, I turned to on-line experiments in which SIs could be measured implicitly.

4. The self-paced reading experiments (chapter 6)

In chapter 5 I discussed a number of on-line studies on the processing of SIs from the literature. Most of these were conducted to settle two debates on the derivation of SIs: the debate whether SIs are calculated after or during compositional semantics (the globalist/localist debate), and the debate whether SIs are generated by default or only when the context licenses them (the defaultist/contextualist debate). I briefly discussed these debates and showed how different types of processing experiments tried to settle them. The goal of discussing these experiments was considering the on-line reflex of SIs that they exposed, which could be used to test the hypothesis of focus-sensitivity of SIs in an on-line experiment. One set of experiments was of particular interest for the current enterprise, as they found effects of the wider context on the processing of scalar terms. These were the reading time experiments of Katsos (2006) (and colleagues). I discussed the findings in one of their experiments about effects of information structure on processing of SIs and showed that they can also be accounted for by the view that SIs are focus-sensitive. However, the crucial finding of Katsos and colleagues that I used as a starting point for the on-line experiments, was that in SI-triggering contexts reading times on the scalar segment were slower than in SI-blocking contexts, presumably due to the processing cost of SI-calculation. Therefore, I decided to conduct an experiment in which I compared reading times on scalar segments in the focus and in the background, using the Self-Paced Reading (SPR) paradigm. If a delay on the scalar region would be observed for the focus-condition compared to the non-focus condition, this would support the hypothesis that more SIs were calculated in the focus condition.

In Experiment 8 I used the minimal focus manipulation of Experiment 2: only the explicit question was varied and no additional material was given in

the story about why the question was asked. In order to control for the effect on reading times of the repetition of the *A or B* region from the question, I included two baseline conditions with *or* replaced by *and*. See (18)-(21) for example items:

(18) Experiment 8 condition 1: focus *or*
Hielke and Sietse were out fishing on their boat.
When **Hielke** returned, their father asked **what kind of fish Sietse had caught**.
Hielke answered that Sietse had caught a pike or a carp.
Father said that Sietse was a lucky devil.

(19) Experiment 8 condition 2: non-focus *or*
Hielke and Sietse were out fishing on their boat.
When **they** returned, their father asked **who of them had caught a pike or a carp**.
Hielke answered that Sietse had caught a pike or a carp.
Father said that Sietse was a lucky devil.

(20) Experiment 8 condition 3: focus *and*
Hielke and Sietse were out fishing on their boat.
When Hielke returned, their father asked what kind of fish Sietse had caught.
Hielke answered that Sietse had caught a pike **and** a carp.
Father said that Sietse was a lucky devil.

(21) Experiment 8 condition 4: non-focus *and*
Hielke and Sietse were out fishing on their boat.
When they returned, their father asked who of them had caught a pike **and** a carp.
Hielke answered that Sietse had caught a pike **and** a carp.
Father said that Sietse was a lucky devil.

As it was expected that the non-focus conditions would be read faster due to the repetition of the *A or B* region, the critical prediction was an interaction between the factors focus (focus vs. non-focus) and connective type (*or* vs. *and*). If more SIs were calculated in the focus *or*-condition than in the non-focus *or*-condition, the difference in reading times on the *A or B* region between these two conditions should be bigger than the difference between the two *and*-conditions. However, the predicted interaction was not observed. The repetition effect was observed: the non-focus conditions were

read faster than their focus counterparts, but this difference was not bigger for *or* than for *and*.

This could mean two things: either participants did calculate more SIs in the focus *or*-condition than in the non-focus *or*-condition, but this did not cause a delay in reading (contra Katsos' findings), or there was no difference in SI-calculation between the focus and the non-focus condition, contrary to the earlier results. An explanation for the lack of an effect of focus was that there was no incentive for the participants to calculate the SI. In the TVJT and PWJT experiments participants had to compare the sentence to the *A and B* situation in order to carry out the experimental task, which made SI-calculation worth the effort. In the SPR-paradigm, they just had to read and memorize the story. As the SI was not relevant for the outcome of the story and the verification statements that were included to check comprehension were never about the SI, there was no reason to calculate it. To avoid this problem I conducted another SPR experiment (Experiment 9) in which I used the richer contexts of the PWJT-experiments. Due to the conditionals included in those contexts, it was relevant for the outcome of the story whether the stronger scalar alternative was true or not, so participants had an incentive to calculate the SI. In Experiment 9 I used the [+alternative relevance, +focus] stories and the [−alternative relevance, −focus] stories of Experiment 5, given in (15) and (16) above. As a result, Experiment 9 tested the cumulative effect of focus and alternative relevance.

In Experiment 9 I introduced another measure of SIs, which was independent of the processing load of SI-calculation itself. I measured reading times on a plural pronoun (*them*) in the next sentence, which referred back to *A or B*. As a plural pronoun is incompatible with the SI-reading of a sentence with *A or B*, but not with the reading without SI, a delay in reading times would indicate that the SI was calculated. Therefore, I predicted a delay would occur after a sentence in which *or* was part of the information focus, but not after a sentence in which it was in the background. As a baseline I used a singular pronoun (*it*), which was compatible with *A or B* irrespective of SI. The penalty should not be observed for the control conditions with *and*, as a plural pronoun is correct for *A and B*.

Example items of the four *or*-conditions are given in (22)-(25). The two control conditions with *and* were identical to condition 1 and condition 3 (as *and* required a plural pronoun), but with all occurrences of *or* replaced by *and*. As the properties of focus and alternative relevance were combined, I used *FocRel* as an abbreviation of the cumulative property.

(22) Experiment 9 condition 1: FocRel or them
Karl told a friend (masc.) that Hugo and Oliver went searching for marine animals on the beach yesterday, and that when they returned, their father said that if one of them had found **at least two marine animals** on the beach, that person got to stay up late that night.
The friend (masc.) said: "And, **what had Hugo found** on the beach?"
Karl answered: "Hugo had found a starfish **or** a mussel on the beach."
Hugo said later that he had found **them** in the surf near the lighthouse.

(23) Experiment 9 condition 2: FocRel or it
Marije told a friend (fem.) that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **at least two marine animals** on the beach, that person got to stay up late that night.
The friend (fem.) said: "And, **what had Laura found** on the beach?"
Marije answered: "Laura had found a crab **or** a mussel on the beach."
Laura said later that she had found **it** in the surf near the lighthouse.

(24) Experiment 9 condition 3: non-FocRel or them
Karl told a friend (masc.) that Hugo and Oliver went searching for marine animals on the beach yesterday, and that when they returned, their father said that if one of them had found **a starfish or a mussel** on the beach, that person got to stay up late that night.
The friend (masc.) said: "And, **who of them had found a starfish or a mussel** on the beach?"
Karl answered: "Hugo had found a starfish **or** a mussel on the beach."
Hugo said later that he had found **them** in the surf near the lighthouse.

(25) Experiment 9 condition 4: non-FocRel or it
Marije told a friend (fem.) that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **a starfish or a mussel** on the beach, that person got to stay up late that night.
The friend (fem.) said: "And, **who of them had found a starfish or a mussel** on the beach?"
Marije answered: "Laura had found a crab **or** a mussel on the beach."
Laura said later that she had found **it** in the surf near the lighthouse.

Contrary to Experiment 8, the data did show the predicted interaction between FocRel and connective type for the reading times on the scalar region. The difference between the FocRel *or*-condition (conditions 1 and 2 combined) and the non-FocRel *or*-condition (conditions 3 and 4 combined)

was bigger than between the FocRel *and*-condition and the non-FocRel *and*-condition, suggesting SI-calculation only happened in the FocRel *or*-condition. Importantly, the plural pronoun *them* in the continuation sentence was also read slower in the FocRel *or*-condition than in the non-FocRel *or*-condition. This delay was not observed when the continuation sentence contained the singular pronoun *it*, nor when the scalar sentence contained *and* instead of *or*. I concluded these results are a strong indication that more SIs were calculated when *or* was in a context in which it was the information focus and in which the stronger scalar alternative is relevant than in a context which lacks these two contextual properties.

We did not observe a delay in Experiment 8 in which only focus was manipulated, but we did observe a delay in Experiment 9 in which alternative relevance was added to the focus condition. Therefore, it is tempting to conclude that focus alone is not enough to trigger SIs. However, we have to be careful in attributing the delay on the scalar region to SI-calculation, as there are alternative explanations for why the delay was observed in the FocRel *or*-condition, and not in the non-FocRel *or*-conditions or the *and* control conditions. The FocRel *or*-condition was the only condition in which the literal meaning was not enough to decide the outcome of the conditional and the only condition in which the requirement of the conditional was not satisfied. Therefore, we cannot rule out that the delay is due to differences in inferential processes about the outcome of the conditional between the FocRel *or*-condition and the other conditions.

The co-occurrence of the delay on the plural pronoun and the delay on the scalar region can also be taken as evidence for the claim that SIs are costly. This is in line with the ideas of the contextualists, who claim SIs are effortful, and goes against the views of the defaultists, who claim SIs are automatic and cheap. However, this conclusion also crucially relies on the assumption that the delay in the scalar region was due to SI-calculation, which is, as I argued above, not certain. Nevertheless, irrespective of this caveat, the data go against a view of default generation of SIs, as no cancellation is reflected in the reading times in the non-FocRel condition, while the lack of a delay on *them* in that condition shows that at that point the SI was absent.

5. The *most*-experiments (chapter 7)

Experiments 1-9 all tested the focus-sensitivity of SIs with items in which the scalar term *or* appeared in direct object position and focus was manipulated by a *wh*-question. In chapter 7 I presented two off-line experiments in which I broadened the scope of the investigation in three ways: First, I tested whether the results could be replicated for another scalar term (*most*) in another scalar position (subject position), by considering question-answer pairs like (26) and (27)

(26) Q: How many of the students drank beer?
A: $\underline{\text{Most}}_{F}$ students drank beer.

(27) Q: What did most students drink?
A: Most students drank beerF.

As in (1) and (2) above, the explicit questions determined that *most* was the information focus in (26), but not in (27), and that therefore the SI that not *all* students drank beer should only be observed in (26). Although in (26) only the scalar term *most* itself was the focus, instead of the whole constituent (as in the *or* experiments), this does not matter for the predictions (see 3.1 of chapter 7). Second, I added a condition in which I tested the predictions of the focus-sensitive accounts in another environment: sentential answers to yes/no-questions, see (28):

(28) Q: Did most students drink beer?
A: Most students drank beer.

The answer in (28) provides no new information (except an implicit 'yes'), hence we can consider it to be focus-less. Therefore, no SIs should arise according to the hypothesis of focus-sensitivity of SIs. Thirdly, I explored the interpretation of scalar terms in yes/no-questions themselves. It is usually claimed that SIs do not arise in questions, and I wanted to explore whether this claim is true. This condition was not directly related to the predictions of focus-sensitivity, although I return to the relation between the two below.

Despite the shortcomings of the TVJT-paradigm which I discussed above, I used this paradigm to test the conditions just presented. The reason for this was that I wanted to be able to compare the results to the off-line results found for *or*, and the TVJT experiments provided the clearest results.

An example story of Experiment 10 is given in (29).

(29) Example story Experiment 10
Five people were present at the get-together of the Celtic language studies program at the University. Several drinks were available.
Sander is a student. He drank beer.
Tom is also a student. He also drank beer.
Eric is a professor. He drank wine.
Martin is a professor too. He drank apple juice.
Frans is a student. He drank beer.

The same story types were used in all conditions. The stories were followed by different dialogues in the different conditions, see (30)-(33):

(30) Experiment 10 Condition 1: (wh focus)
[story in (29)]
A: "How many of the students drank beer?"
B: "Most students drank beer."

(31) Experiment 10 Condition 2: (wh non-focus)
[story in (29)]
A: "What did most students drink?"
B: "Most students drank beer.

(32) Experiment 10 Condition 3 (yes/no sentential answer)
[story in (29)]
A: "Did most students drink beer?"
B: "Most students drank beer."

(33) Experiment 10 Condition 4 (yes/no polar)
[story in (29)]
A: "Did most students drink beer?" yes / no

In conditions 1-3 participants were asked to judge whether the answer given by speaker B was true. A 'false' answer indicated the SI was calculated. In condition 4 participants were asked to answer the question themselves by choosing 'yes' or 'no'. In this condition, a 'no' answer indicated that the question was interpreted as *Did most but not all students drink beer?*, indicating that the SI was calculated.

The SI-rates of experiment 10 were as follows: condition 1: 52%, condition 2: 42%, condition 3: 42%, condition 4: 40%. As the difference between condition 1 and condition 2 was significant, the result that more SIs were calculated if the scalar term was in the information focus was replicated. Condition 3 patterned with the non-focus wh-condition and not

with the focus wh-condition, confirming the predictions based on the view that a sentential answer to a yes/no-question is as a whole background. Finally, the yes/no-question itself did return less SIs than the focus wh-condition, but almost as many SIs as the sentential answer. However, just like in the TVJT-experiments with *or*, the differences were not as black-and-white as predicted by the theory. I explored the option that this is due to the TVJT asking for *truth*. This might have encouraged participants to judge the sentence on its own, ignoring the question it was an answer to. As the question was the critical manipulation, this could explain the small effect. Therefore, I conducted a follow-up experiment (Experiment 11) which was identical except that participants were asked to judge the sentence 'right' or 'wrong' instead of 'true' or 'false'. As right/wrong includes felicity, participants might be more likely to consider the fit of the answer to the question.

The results of Experiment 11 were as follows: condition 1: 64% SIs, condition 2: 48%, condition 3: 44%, condition 4: 38%. Again, the difference between conditions 1 and 2 was significant, replicating the results of focus-sensitivity of *or*. The yes/no-conditions also behaved similarly to Experiment 10. Although the difference between the focus conditions and the other conditions was slightly bigger than in Experiment 11, a between-subjects analysis did not reveal significant differences between the two experiments. Therefore, it seems asking for right/wrong instead of true/false did not make a difference. For completeness I summarize the results of Experiments 10 and 11 in Table 3:

Table 3: Summary SI-rates Experiments 10 and 11

	Condition 1 (wh focus)	Condition 2 (wh n-focus)	Condition 3 (yes/no sent.)	Condition 4 (yes/no polar)
Exp 10	52%	42%	42%	40%
Exp 11	64%	48%	44%	38%

I discussed three main reasons that could explain why the difference was smaller than with *or*. Firstly and most prominently, it is likely that in Experiments 10 and 11 there was carry-over between the conditions. Due to the small item-filler ratio and the similarity of the items, resp. 60% and 55% of the participants gave the same answer on all test items throughout the experiment. The second reason is that the list-like character of the story triggered the implicit double-focus question *Who drank what?* Relative to this question both subject and direct object are part of the information focus, so this could have nullified the manipulation of focus. Thirdly, due to *most* the task involved a lot of counting, which might have distracted away from

the linguistic material and especially the question, which was not relevant for establishing the proportion of e.g. students that drank beer.

Finally, I explored the yes/no-question data in conditions 3 and 4, reconsidering the assumption that answers to yes/no-questions are focus-less. It has been argued in the literature that yes/no-questions might be topicalized, in the sense that one part of it receives extra stress. This type of yes/no-questions is claimed to have a focus-structure of their own. This opens the possibility that the focus-sensitivity of SIs extends to scalar terms in questions.

6. The relation between speaker expertise and the SI of *or* (chapter 8)

In chapters 8 and 9 I further investigated the psychological reality of the claim that the exclusive reading of *or* comes about by an SI. In chapter 8 I discussed a well-known problem that arises when what I called the 'simple Gricean view of SIs' is applied to *or*. This is the view that SIs come about by (some version of) the nonce derivation, repeated here in (34):

(34) i. The speaker used the scalar term *or*.
ii. The speaker could have uttered the same sentence with the scalar term *and* instead of *or*, which would have been stronger / more informative (because the sentence with *and* entails the sentence with *or*).
iii. The sentence with the stronger scalar term *and* would have also been relevant.
iv. The speaker is trying to be as informative as possible (she is obeying the Maxim of Quantity).
v. Apparently, the speaker does not have evidence for the sentence with *and*.
vi. The speaker is well informed.
vii. Therefore, it is likely that the speaker considers the sentence with *and* to be untrue.

The crucial step that is problematic for *or* is step vi. This step is problematic as *or* typically also gives rise to so-called *Clausal* or *Ignorance* Implicatures, the inferences that the speaker does not know the truth value of the disjuncts. In order to derive the SI of *A or B* by the nonce derivation in (34), the hearer has to assume the speaker does not know the truth value of *A* nor the one of *B*, but does know the truth value of *A and B*. I argued this assumption is paradoxical and therefore highly unlikely to be made by the hearer. I dubbed this problem the *Speaker Expertise Paradox*.

In the items of Experiment 4 (which was already partly discussed in chapter 4 – see section 3 above), the experimental question that assessed the SI was preceded by a question about speaker expertise. A complete item of Experiment 4 is exemplified in (35), (the underlined labels were absent in the actual items).

(35) Experiment 4 complete item (condition 2)
Marieke told her mother that Laura went searching for marine animals on the beach yesterday, and that her father had told her that if she would find at least two marine animals, she would get to stay up late that night.
The mother said: "Oh, and what did Laura find?"
Marieke answered: "She found a crab or a mussel."

Question 1 (comprehension-question):
What would Laura be allowed to do if she would find a crab or a mussel?...
Question 2 (speaker-expertise question):
Do you think Marieke knows exactly what Laura found? yes / no
Question 3 (SI-question):
Do you think it is possible that Laura found both a crab and a mussel? yes / no

The second question assessed speaker expertise. A 'no' answer indicated the participant considered the speaker not to be completely informed, which was probably caused by the Ignorance Implicatures that *or* gives rise to. Therefore, I refer to the rate of 'no' answers to Question 2 as the *IgnI*-rate. As every item contained both a speaker expertise question and an SI-question, we could assess whether the problematic situation that participants calculate the SI although they do not consider the speaker to be well-informed, actually occurred. All four conditions of Experiment 4 were tested, (see (9)-(13) above). Of interest was whether it occurred that both Question 2 and Question 3 in (35) above were answered with 'no', irrespective of condition. The results are given in Tables 4 and 5 below. In Table 4 the total percentages of 'no' answers on Question 2 (signaling IgnIs) and Question 3 (signaling SIs) are given per condition.

Table 4: IgnI- and SI-rates Experiment 4

	Condition 1 [-alt +foc]	Condition 2 [+alt +foc]	Condition 3 [-alt -foc]	Condition 4 [control]
IgnI	93%	96%	92%	47%
SI	67%	72%	63%	27%

In Table 4 we see that high IgnI-rates co-occurred with reasonably high SI-rates in all three test conditions. This indicates participants frequently calculated the SI while they did not consider the speaker to be completely informed. This is confirmed when we consider how many times a certain answer combination occurred per condition, see Table 5.

Table 5: Answer combinations given in Experiment 4

	Cond. 1 [-alt +foc]	Cond. 2 [+alt +foc]	Cond. 3 [-alt -foc]	Cond. 4 [control]
Q2: 'no', Q3: 'no' (+IgnI, +SI)	61%	69%	58%	15%
Q2: 'no', Q3: 'yes' (+IgnI, -SI)	32%	27%	34%	32%
Q2: 'yes', Q3: 'no' (-IgnI, +SI)	6%	3%	5%	11%
Q2: 'yes', Q3: 'yes' (-IgnI, -SI)	1%	1%	3%	41%

In the first row of Table 5 we see that in 61% of the times that participants were presented with a condition 1 item, they answered both Question 2 and Question 3 with 'no'. For conditions 2 and 3 this percentage was resp. 69% and 58%. This indicates that in more than half of the cases in each condition, participants did not consider the speaker to be completely informed, yet they interpreted *or* exclusively. I concluded this confirms the Speaker Expertise Paradox and that for that reason it is unlikely the exclusive reading of *or* comes about by an SI derived by the nonce derivation.

In the second part of chapter 8 I presented two types of theories from the literature that might be able to deal with this problem. The first type are theories which hold on to (some version of) the Gricean derivation, but derive SIs in two steps. These are the theories of Sauerland (2004) and Van Rooij & Schulz (2004). They claim the (weak) Clausal / Ignorance Implicatures are derived by Gricean reasoning in the first step, and in the second step the assumption of speaker expertise is added only insofar as it is compatible with the output of the first step (the IgnIs). Although this avoids the problem that adding the assumption of speaker expertise rules out the disjuncts themselves, I argued it does not solve the Speaker Expertise Paradox. The question remains why hearers would take the second step (labeled the *Epistemic Step* by Sauerland and *speaker competence* by Van Rooij & Schulz) of assuming the speaker is informed about the truth value of *A and B*, after they have derived in the first step that the speaker does not know the truth value of A nor of B. I argued we would have to assume the

second step is default. I return to this option in the conclusion section 9 below.

The second type of theory I presented that might avoid the Speaker Expertise Paradox is the view that SIs do not rely on Gricean reasoning. This view is advocated by Fox (2007) and Chierchia (2004, 2006) and Chierchia, Fox and Spector (2008), who claim SIs arise through application of an exhaustivity operator in the grammar. According to Fox (2007), the Ignorance Implicatures (which on his view are derived by Gricean reasoning) might be considered implausible by the hearer. This can trigger a re-parse of the sentence in which an exhaustivity operator is applied to the whole sentence, giving rise to the SI. However, as this re-parse depends on the IgnIs being considered implausible by the hearer, it depends on pragmatic reasoning and therefore the Speaker Expertise Paradox re-arises. Why would the hearer consider the IgnIs implausible for an UE episodic sentence like *Laura found a crab or a starfish*? I concluded Fox's account does not solve the Speaker Expertise Paradox.

Chierchia (2004, 2006) and Chierchia et. al (2008) propose a local exhaustivity operator, which can apply during compositional semantics. However, it is unclear what the contextual/situational criteria for applying this operator are in their model. Chierchia et al. first claim their exhaustivity operator is optional, which gives rise to the Speaker Expertise Paradox. It is unlikely a hearer would apply the operator if the speaker is obviously not well-informed. Then they propose scalar terms have two versions, with and without active alternatives, and selecting the scalar term with active alternatives is optional. This only seems to be moving the problem, as it is unlikely hearers would select the version of *or* with active alternatives in an UE episodic sentence. Finally, they propose that some items only have a strong version which gives rise to obligatory exhaustification. If we assume *or* is one of those special items we do solve the Speaker Expertise Paradox, but we cannot explain the effects of focus observed in the previous chapters. I concluded that in their current form, neither the simple Gricean view, nor the two-step Gricean theories, nor the non-Gricean grammatical exhaustification theories can account for the Speaker Expertise Paradox and the focus-data at the same time. This casts doubt on the view that exclusive-*or* is the result of an SI. I return to this claim in section 9 below.

7. The relation between exhaustivity and the SI of *or* (chapter 9)

In chapter 9 I took another approach to determine whether the SI-view of *or* is correct. I considered the connection between SIs and exhaustivity. One of the theories which predicted the focus-sensitivity of SIs was the theory of

Van Rooij (2002). On his account, SIs arise as the result of the more general procedure of exhaustification of answers. On this view SIs and exhaustivity are the result of the same mechanism, which brings along the interesting prediction that for the answer in (36), the SI in (37) and the exhaustivity inference in (38) (which I dubbed the 'external exhaustivity' inference), should always co-occur:

(36) Q: What did Bart catch?
A: Bart caught a spider or a snake.

(37) Bart did not catch a spider and a snake.

(38) Bart did not catch anything else (besides a spider or a snake).

I argued that not only the theory of Van Rooij (2002) (and the extension of that theory in Van Rooij & Schulz (2004)) makes this prediction, but also the theories which assume a grammatical exhaustivity operator (Fox (2007), Chierchia et al. (2008)). In chapter 9 I showed how this prediction can be derived from their theories.

Experiments 5-7, already partly discussed in chapter 4 (see section 3 above), tested this prediction. Next to the SI-questions asking about the possibility of the *A and B* situation, repeated here from (12) above, the experiment also contained questions about the external exhaustivity inference, exemplified in (39):

(12) Do you think it is possible that Laura had found both a crab and a mussel? yes / no

(39) Do you think it is possible that Laura had also found something else than a crab or a mussel? yes / no

A 'no' answer to (12) indicated the participant calculated the SI. Similarly, a 'no' answer to (39) indicated the participant made the external exhaustivity inference that Laura found nothing else besides a crab or a mussel. As both inferences are claimed to be the result of the same mechanism, they should always co-occur. I chose not to directly assess this by including both questions in every item, but to include one type of question per item and include both types of items in every list. This way it could be avoided that the two questions affected each other in one item. As the two inferences are claimed to be the result of the same mechanism, they should be triggered to the same degree in the same environment. The four conditions of the experiment, in which alternative relevance and focus were manipulated,

were already presented in section 3 above. I repeat the example items here for easy reference.

(14) Experiment 5 condition 1: [–alternative relevance, +focus]
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **a crab or a mussel** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **what had Laura found** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(15) Experiment 5 condition 2: [+alternative relevance, +focus]
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **at least two marine animals** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **what had Laura found** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(16) Experiment 5 condition 3: [–alternative relevance, –focus]
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found **a crab or a mussel** on the beach, that person got to stay up late that night.
The friend said: "Oh, and **who of Laura and Barbara had found a crab or a mussel** on the beach?"
Marieke answered: "Laura had found a crab or a mussel on the beach."

(17) Experiment 5 condition 4: (control)
Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found a crab or a mussel on the beach, that person got to stay up late that night.
The friend said: "Oh, and who of Laura and Barbara had found a crab or a mussel on the beach?"
Marieke answered: "**Laura did.**"

Recall that the theories treat external exhaustivity and SIs as two reflexes of the same mechanism. Therefore, they predict external exhaustivity-rates should be similar to the SI-rates in all three test conditions.

Experiment 5 tested this prediction in a within-subjects design in which different stories were used for the SI and the exhaustivity items. In order to make sure differences were not due to differences between the stories used for the SI-items and the stories used for the exh-items, I used the same stories for both SI-items and exhaustivity-items in Experiment 6, which was set up in a between-subjects design. Finally, I repeated Experiment 5, which was a web-based questionnaire, in a pen-and-paper version (Experiment 7). The results of Experiments 5-7 are summarized in Table 6:

Table 6: Summary SI-rates and external exhaustivity-rates Exp. 5-7

	Condition 1 [−alt +foc]		Condition 2 [+alt +foc]		Condition 3 [−alt −foc]		Condition 4 [control]	
	SI	exh	SI	exh	SI	exh	SI	exh
Exp 5	71%	43%	76%	57%	65%	10%	18%	6%
Exp 6	68%	43%	74%	67%	67%	18%	29%	13%
Exp 7	65%	49%	78%	76%	68%	25%	16%	19%

All three experiments showed a main effect of SI/exh, indicating that across all conditions, participants behaved differently on the SI items than the exh items. All three experiments also showed an interaction between SI/exh and condition, indicating that the manipulations in the conditions had a different effect on the SI-items than the exhaustivity-items. Pairwise comparisons per condition revealed that in Experiment 5 in all three test conditions more SIs were calculated than exhaustifications. In Experiment 6 this was the case in conditions 1 and 3 and in Experiment 7 in condition 3. However, in Experiments 5 and 6 the difference in the control condition was also significant. If we control for this baseline difference by considering the interactions (of SI/exh and test/control), in two out of three test conditions in Experiment 5 (C1 and C3) and one of three in Experiments 6 and 7 (C3), significantly more SIs were calculated than external exhaustivity-inferences. Individual differences between participants or stories cannot have been responsible for the difference, as the same participants answered SI- and exh-questions in Experiments 5 and 7 and the same stories were used for SI- and exh-items in Experiment 6. The results of Experiments 5-7 suggest that SIs and external exhaustification do not come about by the same mechanism, which goes against the theories discussed above. This raises more doubts about the claim that exclusive-*or* comes about by an SI.

However, as I already pointed out we have to be careful in interpreting the results. When I discussed the SI-data of Experiments 4-7 in chapter 4, I

suggested that the effect of the critical manipulations was overruled by the experimental question (see section 3 above and section 7.2 of chapter 4 for the argument). The reading time data of Experiment 9 confirm this. In that experiment, the same story-types were used as in Experiments 5-7 and there was a clear difference between what are here condition 2 and condition 3. However, as the results in Table 6 indicate, the exhaustivity data seem to be less susceptible to this masking effect. I argued this might be due to the fact that in the SI-questions, there was a contrast between overt material (*A or B*) in the target sentence and the questioned situation (*A and B*). This contrast was absent in the exhaustivity items. In those items, the question was about something else next to A or B, and the target sentence only contained *A or B*. This might have forced participants to look at the wider context for clues in the exhaustivity-conditions, while in the SI-conditions they based their answer on the contrast with the target sentence. The fact that the exhaustivity-questions contained the phrase *also something else than A or B*, might have also encouraged participants to look beyond the target sentence. As the wider context contained the critical manipulations, it is not surprising these had a bigger effect in the exhaustivity-items. This could explain why the masking effect of the experimental question was bigger for the SI-questions.

However, from the perspective of the exhaustivity theories, it is not so obvious that the two questions should lead to different behavior. Only on the exclusive reading of *or* it is incompatible with the questioned situation. The contrast with the questioned situation is therefore only present if the exhaustivity operator is applied (or if a participant recognized the fact that applying it would lead to a contrast). But application of the exhaustivity operator also gives rise to a contrast between the questioned situation in the exhaustivity-items (*A or B and X*) and the target sentence (*A or B and nothing else*). On this view it remains strange that the two types of questions led to such different behavior. Although the possibility of interfering methodological factors make it hard to draw conclusions, I argued the data at least raise doubts about the view that SIs and exhaustivity are the result of the same mechanism. As many influential SI-theories rely on (some form of) exhaustivity, this is another reason to doubt the SI-view is correct for *or*.

8. Conclusions and suggestions for further research on the context-sensitivity of SIs

In this dissertation I addressed the question what the contextual properties are that determine whether SIs arise. I adopted the predictions of Van Kuppevelt (1996) and Van Rooij (2002) that SIs only arise if the scalar term

is in the information focus relative to the contextual question and tested it in a series of experiments in different experimental paradigms. The experimental results supported the hypothesis that SIs are sensitive to focus. The TVJT-experiments showed that for one and the same sentence such as *Harry brought bread or chips*, more SIs were calculated when the scalar term *or* was in the information focus than when it was in the background. This effect was observed for target sentences which were answers to explicit contextual questions, as well as to implicit contextual questions in which the information focus part was marked by stress. The fact that the effect was replicated for another scalar term (*most*) in another structural position further supports the hypothesis and suggests that it is indeed a general property which holds for all SI-triggers. Furthermore, the prediction of the focus-sensitive view that fewer SIs are calculated if a scalar term is in a sentential answer to a yes/no-question (which is as a whole background) than if it is in the information focus, was also confirmed. I conclude that the focus is indeed a contextual property that determines whether SIs arise.

In a number of experiments, the effect of focus only surfaced when it was combined with another contextual property which is claimed to be crucial for SIs, and which is hard to tease apart from focus: the property of relevance of the stronger alternative to the hearer ('alternative relevance'). I implemented this property through the use of explicit conditionals in the context. By manipulating the requirement in the conditionals, the stronger alternative could be made relevant or irrelevant. I found support for the effect on SI-calculation of the cumulative property of focus and alternative relevance in off-line as well as on-line experiments. In two of the four experiments conducted in a new off-line experimental paradigm called the Possible World Judgment Task, more SIs were calculated when the scalar term was in the information focus and the stronger scalar alternative was relevant than when these properties were absent. In a self-paced reading experiment the plural pronoun *them* in a continuation sentence was read slower after a sentence in which *or* was in the information focus and a prior conditional had triggered alternative relevance, than when these properties were absent. As a plural pronoun is incompatible with an SI-reading of *A or B*, the penalty indicates that more SIs were calculated in the condition with focus and alternative relevance than in the condition with neither. The hypothesis that the delay was due to an SI was supported by the absence of a delay for the control conditions with *and* and the fact that no difference was observed when the continuation pronoun was singular (*it*). Finally, the condition with focus and alternative relevance also showed a delay in reading on the scalar region itself compared to the condition with neither properties. This also supports the view that more SIs were calculated in the focus+alternative relevance condition, as independent results in the literature

have suggested SI-calculation is costly. These results show that focus and alternative relevance together increase SIs. I will now turn to the question of their mutual relation.

I acknowledged that focus and alternative relevance are hard to tease apart. Therefore, it is also hard to draw conclusions about which one of them is more important than the other. Leaving aside the objections against drawing conclusions based on a null-result, it is tempting to conclude from a comparison of the first and the second SPR experiments (Experiments 8 and 9), that alternative relevance is the crucial property. In Experiment 8 I only manipulated focus, and no delay on the scalar region was observed in the focus condition. In Experiment 9 I added alternative relevance to the focus condition, by including a conditional for which the stronger scalar alternative was relevant, and the delay was observed. I argued this difference was due to the lack of an incentive for the participant to calculate the SI in Experiment 8. As in Experiment 9 the conditional triggered this incentive, more SIs were calculated in the focus-condition. However, an alternative view is that no effect was observed in Experiment 8 because focus is not the crucial property, but alternative relevance is. On this view the effect in Experiment 9 was due to the added alternative relevance difference.

I pointed out two reasons why we cannot draw the conclusion that alternative relevance is the crucial property. The first is that we cannot be certain that the delay in the scalar region in Experiment 9 was due to SI-calculation. The focus+alternative relevance condition in Experiment 9 was the only one in which the lexical meaning of the connective (*or/and*) was not enough to determine the outcome of the conditional. Also, it was the only condition in which the requirement in the conditional was not satisfied. Therefore, inferential processing about the outcome of the conditional could also have been responsible for the delay. The second reason we cannot conclude alternative relevance is more important stems from a comparison of (TVJT) Experiment 2 and Experiment 8, in which the same stories were used. If alternative relevance was responsible for the difference in the (TVJT) Experiment 2, we have to assume the focus-question and the non-focus question introduced a difference in alternative relevance, as only the questions varied between conditions. However, Experiment 8 also contained this difference, so then it is puzzling why no effect was observed there. We would have to resort the the hypothesis that contrary to Experiment 2 there was no incentive for SI-calculation in Experiment 8. But then we are back where we started and the difference between Experiments 8 and 9 tells us nothing about which property (focus or alternative relevance) is the crucial one anymore. In chapter 6 I proposed testing a [-alternative relevance, +focus] condition in a follow-up SPR experiment which might get us out of this circle. I leave this as a suggestion for further research.

However, the main problem in teasing apart focus and alternative relevance is that alternative relevance is not well-defined. I claimed it might be triggered by a question, and I implemented it in the experiments with an explicit conditional, but this obviously is not the only way in which the stronger scalar alternative can be made relevant. In a sense, the property of alternative relevance suffers from the same problem that I claimed many pragmatic theories suffer from: it is too vague a notion to make clear predictions. Therefore, I argue the next step in defining the contextual properties which determine SI-calculation should be to make notions like these more explicit. For now I stick to my interpretation of the data that the experiments provided experimental support for the effect of a contextual property which is well-defined: information focus.

9. Conclusions and suggestions for further research on the methodology for assessing SIs

The methodological goal of this thesis was to find the most suitable paradigm to test the hypothesis of the focus-sensitivity of SIs and to assess SIs in general. In section 6 of chapter 3 I discussed the TVJT and I argued it is not very suitable for assessing SIs, because it requires the actual situation to be given. This goes against the way in which SIs arise 'in the wild', where hearers do not know what happened and draw conclusions based on what the speaker said. I argued it is highly unlikely that in the TVJT participants interpret the sentence as if they did not know what happened, calculating the SI. Rather, participants simply check whether the sentence fits with the story. Therefore, a 'false' answer is more likely to be based on the use of an underinformative scalar item than on an SI. Even if the linguistic contexts in which participants are most prone to reject underinformative items are the same contexts in which they would calculate SIs if they did not know what happened, the TVJT is still a very indirect measure. In the same section I pointed out that another problem of the TVJT is that it asks for a judgment based on an interpretation instead of (a part of) the interpretation itself. This creates problems with sentences which seem to be 'partly true'. If a more direct measure of the interpretation is asked for (as in the PWJT), these sentences are not problematic. Furthermore, as I pointed out in chapter 7, the focus on truth of the target sentence might encourage participants to ignore contextual properties and focus on the sentence itself. This would make the TVJT very unsuitable for assessing contextual effects. However, I did not find a difference when I asked participants for right/wrong judgments instead of true/false judgments in Experiments 10 and 11.

The new paradigm that I presented in chapter 4 solved many problems of the TVJT. In the PWJT the situation was no longer given and therefore it was a much more natural environment for SIs to arise. Furthermore, the crucial part of the interpretation was assessed directly, by asking about the possibility of the *A and B* situation. This allowed us to target the part of the interpretation that was relevant for our purposes, solving the problem of 'partly true' answers. However, this directness was also the weak point of the paradigm, as I pointed out in section 7.2 of chapter 4 (and in section 3 above). As the experimental question explicitly asked about the situation which could have been described by the stronger scalar term, it made the issue of whether the stronger scalar term held relevant (see Geurts in prep.), overruling the manipulations of alternative relevance. Also, it asked about the direct object in all conditions, overruling the focus-manipulations. As the experimental question was presented after the story, it was the most recent issue that was relevant when participants went back to the target sentence to answer the question. This is contrary to the TVJT, where the explicit question which contained the critical manipulation was the most recent issue. This might explain why the effect of the critical manipulation was observed in the TVJT, but was smaller or absent in the PWJT. The hypothesis that the small differences in the PWJT-experiments are due to the experimental question is confirmed by the fact that we found clear effects with the same type of items in the self-paced reading Experiment 9. The delay on *them* indicated that more SIs were calculated in the focus condition, while this effect was small or absent in the PWJT-experiments, in which the same items were used. I conclude that in order to assess SIs we have to avoid asking about the critical situation explicitly, and we have to make sure the critical manipulation is the most relevant issue when the target sentence is interpreted.

Therefore, a more suitable type of paradigm to assess SIs are on-line experiments, as these asses interpretation implicitly. This allows us to leave out the actual situation so SIs can arise naturally (contra the TVJT), At the same time, we can leave out the experimental question which affects participants' behavior (contra the PWJT). In Experiment 9 I used two measures of SI: a delay in reading of the scalar region itself and a delay in reading on a region in the next sentence which was incompatible with the SI. I argue the second measure is the best way to assess whether participants calculated an SI or not. Although there is a growing number of studies that claim to have found a processing cost of SIs, it is hard to rule out other explanations for a delay on the scalar region itself. Overall, reading time studies are a suitable paradigm to measure contextual effects on SIs and to assess SIs in general. The disadvantage is that it does not provide absolute data on the proportion of items in which participants calculated the SI.

However, we did run into another limitation of using reading times to assess SIs. In Experiment 8 no delay on the scalar region was observed in the focus-condition. I argued this was due to the fact that there was no incentive for participants to calculate the SI. Although we cannot conclude anything from a null-result, let us for the sake of the argument assume that indeed no SIs were calculated in the focus condition in Experiment 8. We might then conclude that focus is not a sufficient condition for SI-calculation. However, this is not necessarily the case. Perhaps if the participants would have been in a dialogue similar to the ones in the stories themselves, they would have calculated the SIs. If participants themselves are in a situation in which they are trying to find out what happened instead of a character in a story, they might be more likely to make pragmatic inferences such as SIs. Therefore, an interesting direction to take for experimental research in pragmatics would be to conduct more experiments in which the participant herself is the addressee and she has to find out what happened. Interesting work in this direction has been done by Mol et al. (2004), who investigated pragmatic interpretations of participants who were involved in a game of Master Mind, in which one has to guess the opponents' secret code. This kind of experiment is very suitable for assessing pragmatic interpretations such as SIs in a way that closely resembles the way they arise in everyday conversation, but yet allows us to control and assess the conditions under which they are derived.

10. Conclusions and suggestions for further research on the psychological reality of the SI-view of *or*

Throughout this thesis, I have investigated the psychological reality of SI-theory, and of the SI-view of *or* in particular, by testing under which circumstances SIs arise and testing their co-occurrence with other inferences. There are at least four topics on which I provided experimental data in this thesis that seem to be relevant for theories of how SIs, and the SI of *or* in particular, are derived. The first one is the focus-sensitivity of SIs. In Experiments 1-9 a clear pattern emerged that more SIs are calculated if *or* is part of the information focus, relative to a preceding question (see sections 2-5 above for an overview. The second issue concerns the processing cost of SIs. Experiment 9 showed that in the focus condition the region containing the scalar term *or* was read significantly slower than the same region with *and*, while no such delay was observed in the non-focus condition. As *and* cannot trigger SIs (in UE sentences) and *or* can, and we found more SIs for the focus-condition in off-line experiments, this suggests the delay was due to SI-calculation. This was supported by the delay on the subsequent plural

pronoun in the same condition, which was absent for the non-focus condition and the control conditions with *and* and *it*. On the face of it, this provides strong evidence for the view that the delay we observed on the scalar region was due to SI-calculation, and therefore SIs are costly. Although I argued we have to be careful in attributing the delay on the scalar region to SI-calculation, the experiment at least showed no cost of cancellation in the non-focus condition, while the absence of a delay on the plural pronoun in the next sentence indicated the SI was absent at that point.

The third issue is the Speaker Expertise Paradox. Experiment 4 showed that participants calculated the SI of *or* even though they did not consider the speaker to be completely informed, which is a well-known problem for the standard Gricean account. Finally, the fourth issue is the relation between SIs and exhaustivity. Experiments 5-7 showed that the SI of *A or B* (n*ot A and B*) was more frequent that the external exhaustivity inference (*not C*). Although I pointed out these data might be affected by methodological issues, I argued this was not so obvious from a the point of view of the exhaustivity theories.

I discussed four types of SI-theory in this thesis, which I argue cover a large part of the dominant views in the literature at the moment. The first is the view which I called the 'simple Gricean view' in chapter 8: the view that SIs are derived by (some version of) Gricean reasoning along the lines of the nonce derivation. These are the views that consider SIs to be Particularized Conversational Implicatures (PCIs) that are dependent on the particular situation and context. I consider Relevance Theory to also belong to this camp. Although they replaced Grice's view with their own principles, they consider SIs to be situation- and context-dependent and costly. For clarity, I will call this view the *PCI-view* from now on. The second view is the view which I described as the 'defaultist' view in chapter 5. This is the view that the SI arises automatically when a scalar term is encountered, to be cancelled if the context does not license it. On this view, most clearly advocated by Levinson (2000), generation of SIs is cheap but cancellation might be expensive. The third view is the Gricean exhaustivity view of Van Rooij & Schulz (2004), discussed in chapter 8. Their account is Gricean in the sense that their exhaustivity operator formalizes Gricean reasoning: it is a pragmatic procedure which applies to the output of compositional semantics and is based on the maxims of Quality and Quantity. The fourth view is the grammatical view of SIs advocated by Chierchia and colleagues. This view proposes SIs come about by application of an exhaustivity operator in the grammar. This operator can be applied locally, during compositional semantics. I argue none of the four theory-types mentioned here can account for the whole range of data presented in this thesis. Let us consider the four

theory-types one by one and consider their fit to the four issues mentioned above.

The PCI-view does not explicitly predict the focus-effect, but it can account for it by relying on the vague notion of (alternative) relevance. The processing findings were also in line with this view: SI-calculation requires contextual support and is costly. As contextual support was only present in the focus condition and as SI-calculation is costly, this led to the delay in the focus condition. However, a hard problem for this view is the Speaker Expertise Paradox. I argue it is unlikely that a hearer will assume/derive that the speaker does not know the truth value of the disjuncts, and at the same time will make the assumption that the speaker is well-informed about the truth value of *A and B*. Finally, the exhaustivity data are no problem for the PCI-view, as on this view SIs do not come about by application of an exhaustivity operator.

The defaultist view does not predict the focus-sensitivity of SIs, but it can account for it by relying on cancellation of the SI in the non-focus condition. However, this view cannot account for the processing findings. Firstly, it cannot explain the delay in the focus condition. As SI-calculation is automatic and cheap, it should not increase processing load. Also, if SIs despite being cheap would lead to a processing load, this should also be observed in the non-focus condition, as generation is default. Furthermore, in that condition no delay was observed on *them* in the next sentence, suggesting the SI was absent at that point. Therefore, an effect of cancellation should have been observed somewhere in the region from the scalar term to *them*. We found no indications of such a process. The Speaker Expertise Paradox is no problem for the defaultist view as generation does not rely on speaker expertise. However, it can only avoid the SEP if SIs cannot be cancelled because of obvious epistemic limitations of the speaker. If they can, the question arises why SIs would ever survive for *or* in UE episodic sentences. Finally, the exhaustivity data are not relevant for the defaultist view, as SIs and exhaustivity are not the result of the same mechanism.

The Gricean exhaustivity view of Van Rooij & Schulz (2004) predicted the effect of focus, so the data in chapters 3-7 support their theory. They can also account for the processing data, on the assumption that application of the exhaustivity operator(s) is costly. As the operator only exhaustifies the material in the focus, the delay was only observed in the focus-condition. This also explains why the subject region was read slower in the non-focus

condition, in which it was the focus region.[1] The fact that the delay was observed at the scalar region itself seems to go against the claim that the exhaustivity operator is applied after compositional semantics, but many authors have pointed out that post-compositional calculation of SIs does not exclude that participants start calculating them before the sentence ends, for instance based on expectations about how the sentence is going to end. For instance, Geurts (in prep.) says:

> 'While it is true that a full-fledged conversational implicature requires a full-fledged sentence meaning, this doesn't mean that hearers can't process implicatures on the fly, just as there is no reason why entailments can't be computed incrementally.' (p. 73)

As the scalar region in Experiment 9 was close to the sentence end and the rest of the sentence was highly predictable as it was repeated from the question, these data do not go against post-compositional calculation of SIs.

However, the account of Van Rooij & Schulz cannot account for the Speaker Expertise Paradox. They propose a two-step mechanism (with two exhaustivity operators) for deriving the SI. In the first step the Clausal Implicatures are derived and in the second step the assumption that the speaker is well-informed is added and the SI is derived. Van Rooij & Schulz say:

> 'Hence, we predict the strong reading [of the SI, AZ] to occur in those contexts where such a competence assumption can be made. Otherwise only the weak reading is observed.' (p. 27)

However, I argue it is unlikely hearer will make such an assumption in UE episodic sentences after the Clausal Implicatures have been derived in the first step. I argued the only way around this problem is to assume the second step is taken by default, irrespective of considerations of speaker expertise. I return to this option below. Furthermore, the exhaustivity data presented in chapter 9 suggest SIs and external exhaustivity are not the result of one mechanism, which is problematic for Van Rooij & Schulz (although caution is in order as the methodology might have added to the difference).

Finally, let us consider the grammatical exhaustivity account of Chierchia and colleagues. This account is difficult to assess as the situational/contextual criteria for application of their exhaustivity operator

[1] It does however not explain the unexpected finding that the subject region was read slower in the non-focus *or* condition than in the non-focus *and* condition (see section 3.5.2 of chapter 6). I come back to this difference later.

have not (yet) been specified. Let us assume the application is optional (see chapter 8 section 5.2.2 for quotes by Chierchia and colleagues in which this suggestion is made). Notice that the Strength Condition of Chierchia (2004), which is built into the exhaustivity operator of Chierchia (2006) cannot account for the focus-difference in UE episodic sentences, and contextual information (such as the preceding question) has to be taken into account. As I argued in chapter 9 section 2.4, it would be nice from a conceptual point of view if the scalar alternative set and the focus alternative set are somehow related, and Fox & Spector (2008) seem to make this claim. This could possibly account for the focus-sensitivity of SIs, although it remains to be specified how it would follow from their account. In order to account for the processing data on Chierchia et al.'s view, we have to assume application of the exhaustivity operator is costly (and absent in the non-focus condition). However, if we assume application of the exhaustivity operator is optional, or (as Chierchia (2006) suggests) the speaker selects the version of the scalar term with or without active alternatives depending on the context, we run into the SEP. Why would the hearer ever select the version of *or* with active alternatives in an UE episodic sentence? Finally, I argued Fox & Spector (2008) make the claim that the SI of *or* should co-occur with external exhaustivity, which was not observed in Experiments 5-7. We could propose a version of Chierchia's account in which *or* triggers obligatory exhaustification, but this version cannot account for the focus-sensitivity.

Summing up, all four types of SI-theory discussed above seem to be unable to account for the full range of data presented in this thesis. The main problem is that we need an SI-theory that is context-sensitive enough to account for the focus-sensitivity, yet default enough to avoid the Speaker Expertise Paradox. As I said above, we might be able to achieve this by assuming the second step of the theory of Van Rooij & Schulz is made by default. On this view, the focus-sensitivity is still accounted for, as the exhaustivity operator still only applies to the focus, and we avoid the SEP as we do not need speaker expertise for the SI to arise. Notice that making the second step of the exhaustification procedure default does not completely reduce the theory to the defaultist theory discussed above. Contrary to the defaultist theory, no SIs are calculated in non-focus environments. Therefore, the processing data can also be explained, on the assumption that the total exhaustification procedure is costly. However, the exhaustivity data remain problematic and we would have to claim that these are completely due to side-effects of the experimental paradigm. Furthermore, the assumption that the second step is default makes the strong prediction that Clausal / Ignorance Implicatures and (strong) Scalar Implicatures should always co-occur. As I pointed out in chapter 8, this prediction is not contradicted by the data of Experiment 4. Although the IgnI-rates were

higher than the SI-rates, the difference was in the same range as the difference in the baseline condition 4). I leave this strong prediction as a suggestion for further research.

The claim that none of the four SI-theories presented above can account for the range of data in this thesis, raises serious doubts about the SI-view as the right account of the apparent ambiguity of *or* between an inclusive and an exclusive reading. Therefore, another possible direction to take would be to abandon the traditional view that inclusive-*or* is the basic meaning of *or* and that exclusive-*or* is derived by an SI. The conceptual argument in favor of the traditional view of Grice and his followers was that we can keep one lexical meaning of *or*, which is to be preferred over assuming an (unnecessary) lexical ambiguity (Grice's 'Modified Occam's Razor'). However, we could also derive the two readings in the opposite direction: from a basic exclusive-*or* to a derived inclusive-*or*.

A proposal along these lines was made by Kratzer (2003). Her account is set in the framework of situational semantics (see Kratzer 2009). I will not go into the implementation of her account in this framework in detail, but I will try to get the intuition across. Kratzer proposes there are two types of assertions: Strong assertions which say the proposition is true in the actual situation and weak assertions which say that there is an actual situation in which the proposition is true. If we take the exclusive meaning of *or* (ex-or) as basic, the strong assertion is that *A ex-or B* is true in the actual situation. The weak assertion is that there is a situation in which *A ex-or B* is true. Due to the existential quantification, this latter assertion comes out true in a situation in which *A and B* is true. In order for there to be a situation in which *A and B* is true, there has to be an actual situation in which either *A* or *B* is true. Hence, whenever *A and B* is true, the weak assertion is true. So we derive the apparent ambiguity of *or* by taking ex-*or* as basic and weakening it by an existential quantification.

Kratzer states that out of context, the strong assertion will be preferred, since it is more informative to state that a proposition is true in the world as a whole or in a salient actual situation than to state that there is an actual situation in which the proposition is true: the first entails the second. However, context (e.g. certain questions) can overrule this preference. Kratzer provides examples with numerals. These are traditionally also considered to be scalar terms, but Kratzer takes their basic meaning to be the *exactly*-meaning , see (40):

(40) A: How many towels did you use?
B: # Two. In fact, I even used three.

(41) A: Did anybody use two towels?
B: Yes, I did. In fact, I even used three.

In (40) the question is a request for a strong assertion. Speaker A wants to know how many towels speaker B used in the actual world. Therefore, the answer of speaker B is a strong assertion, and as the lexical meaning is taken to be *exactly two*, the *In fact...* sentence is infelicitous. However, in (41) what is at issue is an actual situation in which someone used two towels. This is a request for a weak assertion. Therefore, speaker B can answer positively: As she used three towels in the actual situation, there is an actual (sub)situation in which she used two towels.

The relation with the focus-data in this thesis becomes clear from comparing our question-answer pairs (1) and (2) to (40) and (41).

(1) Q: What did Harry bring?
A: Harry brought bread or chips$_F$.

(2) Q: Who brought bread or chips?
A: Harry$_F$ brought bread or chips.

In Kratzer's terms, the question in the focus condition (1) is a request for a strong assertion, like (40). Therefore, the answer is taken to be a strong assertion and *or* is interpreted on its basic ex-or meaning. However, the question in the non-focus condition (2) is a request for a weak assertion, like (41). It can then be paraphrased as *For who does it hold that there is a situation in which he/she has A ex-or B?* The answer would then be the weak assertion *For John it holds that there is a situation in which he has A ex-or B*, which has the truth conditions of inclusive-*or*. Kratzer also proposes an explanation for the preference for the inclusive reading in DE-environments. She introduces two types of negation, corresponding to the two types of assertions: A *weak denial* is the negation of a strong assertion and a *strong denial* is the negation of a weak assertion. Let us see how this works out for *or*, see (42):

(42) John didn't catch A or B.

The weak denial reading is that in the actual world it is not the case that John caught either A or B. This is the reading which is often claimed to involve metalinguistic negation (see Horn (1989)). On Kratzer's account, it is normal (logical) negation. This reading allows for the situation that John caught both. The strong denial reading is that there is not an actual situation in which John caught either A or B. That means that there cannot be a situation

in which John caught both, as whenever there is a situation in which John caught both, there is always a subsituation in which he caught one of the two. This is what Kratzer calls 'the magic of strong denial': the exclusivity of *or* disappears on its own. Nevertheless, the inclusive reading that results is derived from the basic exclusive reading.

Kratzer's implementation in situation semantics is just one implementation of the idea of a basic exclusive meaning and a weakening by existential quantification. We could also propose an existential quantification over flexible types or over propositions (see e.g. Aloni (2007)). For instance if the weakened version of (42) is *there is not a thing which John caught which is either A or B*, the 'magic of strong denial' that he cannot have caught A and B also follows (because if he caught A and B there is also something that he caught which is either A or B). On this implementation the questions in the non-focus condition *Who caught A or B?* would mean something like *For who does it hold that there is a thing which he caught which is either A or B?* which seems intuitively right.

Let us consider how well a view like this accounts for the data. I already illustrated it can account for the focus-effect. The processing findings however seem problematic. If exclusive-*or* is basic and inclusive-*or* is derived by existential quantification, we should have observed a delay in the non-focus condition instead of the focus condition. However, if we assume the existential quantification was already present in the question, no delay is expected in the non-focus condition either. Anyway, on this view the LFs of the target sentences were completely different in the focus condition and in the non-focus condition. As the latter contained the existential quantification, they were not comparable to begin with. Furthermore, as I argued in chapter 6, the delay on the scalar region could be due to inferential processing about the outcome of the conditional. The view that exclusive-*or* is basic could also shed light on the unexpected result that the subject of the target sentence was read slower in the non-focus *or*-condition than in the non-focus *and*-condition. On the exclusive-*or*-as-basic view, the LFs of the questions *Who has A and B?* and *Who has A or B?* were completely different, see (42) and (43), where $\veebar$ is the symbol for exclusive-*or*.

(42) $?\lambda x(HAS(x, A \oplus B)$

(43) $?\lambda x \exists y[(HAS(x,y) \wedge (y = A \veebar y = B)]$

Perhaps it gives rise to a higher processing load to fill in the value of the variable x in (43) than in (42), leading to the longer reading times on the subject in the non-focus *or* condition. However, at this point this is only speculation so I leave this for further research. In general, I argue the

processing findings in the target sentence are not necessarily incompatible with a view in which exclusive-*or* is basic and inclusive-*or* is derived through existential quantification. The main difference with the defaultist view, for which the processing findings are problematic, is that no cancellation effect is predicted in the non-focus condition.

An important advantage of the exclusive-*or*-as-basic view is that we avoid the Speaker Expertise Paradox. As exclusive-*or* is the basic meaning, we do not have to rely on reasoning about speaker expertise anymore. The Clausal Implicatures can stay as they are. Finally, the fact that exclusive-*or* seems to be less context-sensitive than exhaustivity inferences is also no longer a problem. Notice that the exclusive-*or*-as-basic view can still explain why the rates of exclusive interpretations in the focus conditions were lower than 100% in the experiments. For instance in the TVJT, there was a reading which made the sentence true: the weakened existential quantification reading. So although the context (in the focus condition) triggered the stronger basic exclusive meaning, charitable participants might have taken the weaker meaning as it made the sentence true.

A final argument in favor of an exclusive-basic view like Kratzer's is anecdotal. Many, if not all of the people I talked to about this research who had no knowledge of the topic think the meaning of *or* is exclusive. They usually say it means 'one of the two'. One of the ways in which I have tried to explain that *or* at least sometimes is inclusive, is by presenting them with a sentence in which a disjunction is embedded under negation, such as *It is not the case that John has A or B*. I then point out that not only the situations in which John has one of the two are negated, but the situation that he has both is also negated. In other words, the sentence does not mean *It is not the case that John has one of the two*, as that would allow for the situation in which John has both. However, my interlocutors are usually unimpressed. They typically shrug and say: "Of course, if John does not have one of the two, he also cannot have both", followed by a triumphant: "But that does not mean *or* means *and/or*!". It seems they very cleverly switched from wide scope negation to narrow scope negation. However, on the exclusive-*or*-as-basic account, this remark expresses exactly 'the magic of strong denial'. On this view, the speaker was actually saying: "Of course, if there is not a thing which John has which is A or which is B, he also cannot have both." This illustrates that the exclusive-*or*-as-basic view fits well with naive intuitions. Therefore, I end this dissertation by suggesting a further investigation of the hypothesis that exclusive-*or* is basic.

BIBLIOGRAPHY

Aloni, M. (2007). Free choice, modals, and imperatives. *Natural Language Semantics* 15: 65-94.

Ariel, M. (2004). Most. *Language* 80: 658-706.

Banga A., I. Heutinck, S.M. Berends & P. Hendriks (2009). Some implicatures reveal semantic differences. In: B. Botma & J. van Kampen (eds.) *Linguistics in the Netherlands 2009*: 1-13. John Benjamins, Amsterdam.

Bezuidenhout, A. & J. C. Cutting (2002). Literal meaning, minimal propositions, and pragmatic processing. *Journal of Pragmatics* 34: 433-456.

Bezuidenhout, A. & R.K. Morris (2004). Implicature, relevance and default pragmatic inference. In: I. Noveck & D. Sperber (eds.) *Experimental pragmatics*: 257–283. New York: Palgrave Macmillan.

Bezuidenhout, A., R. Morris and C. Widmann (2009). The DE-blocking hypothesis: the role of grammar in scalar reasoning. In: U. Sauerland & K. Yatsushiro (eds.) *Semantics and Pragmatics: From Experiment to Theory*: 124-144. New York: Palgrave Macmillan.

Birch, S. & K. Rayner (1997). Linguistic focus affects eye movements during reading. *Memory & Cognition* 25 (5), 653-660.

Blutner, R. & R. Sommer (1988). Sentence processing and lexical access: The influence of the focus-identifying task. *Journal of Memory and Language* 27: 359-367.

Bott, L. & I. Noveck (2004) Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language* 51(3): 437–457.

Bredart, S. & K. Modolo (1988) Moses strikes again: focalization effect on a semantic illusion. Acta Psychologica 67:135-144.

Breheny, R. (2005) Some scalar implicatures really aren't quantity implicatures – but some's are. In: E. Maier, C. Bary and J. Huitink (eds.) *Proceedings of Sinn und Bedeutung 9*. Nijmegen: Nijmegen Centre of Semantics.

Campbell, R. N. (1981). Language Acquisition, Psychological Dualism and the Definition of Pragmatics. In: H. Parret, M. Sbisa, and J. Verschueren (eds.) *Possibilities and Limitations of Pragmatics*. John Benjamins, Amsterdam.

Carlson, K., M. Walsh Dickey, L. Frazier and C. Clifton (2009). Information structure expectations in sentence comprehension. *Quarterly Journal of Experimental Psychology*. 62(1): 114–139.

Carston, R. (1998). Informativeness, relevance, and scalar implicature. In R. Carston & S. Uchida (eds.) *Relevance theory: applications and implications*. Amsterdam: Benjamins.

Chemla, E. and B. Spector (2010). Experimental evidence for embedded implicatures. In: M. Aloni and K. Schulz (eds.) *Proceedings of the Amsterdam Colloquium 2009*. Springer.

Chevallier, C., I. Noveck, T. Nazir, L. Bott, V. Lanzetti, & D. Sperber (2008). Making disjunctions exclusive. *The Quarterly Journal of Experimental Psychology*, 61(11), 1741–1760.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena and the syntax/pragmatic interface. In: A. Beletti (ed.) *Structures and Beyond*. Oxford University Press.

Chierchia, G. (2006). Broaden Your Views: Implicatures of Domain Widening and the 'Logicality' of Language. *Linguistic Inquiry* 37(4), 535–590.

Chierchia, G., S. Crain, M.T. Guasti, A. Gualmini and L. Meroni (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In A. H.-J. Do et al. (eds.) *BUCLD 25 Proceedings*,157-168.

Chierchia, G., B. Spector, & D. Fox (2008). The Grammatical View of Scalar Implicatures and the Relationship between Semantics and Pragmatics, In P. Portner, C. Maienborn & K. von Heusinger, K. (eds.) *Handbook of Semantics*, Mouton de Gruyter.

Chierchia, G., L. Frazier & C. Clifton (in press). When basic meanings are (not) enough: Processing scalar implicatures in adult comprehension. M. Biezma, C. Davis & J. Harris (eds.) *UMOP 39: Papers in Pragmatics*. Amherst, MA: GLSA Publications.

Chomsky, N. (1971). Deep Structure, Surface Structure, and Semantic Interpretation. In: D. Steinberg and L. Jakobovits (eds). *Semantics: An Interdisciplinary Reader*: 183–216. Cambridge: Cambridge University Press.

Chomsky, N., & M. Halle (1968). *The sound pattern of English*. New York: Harper & Row.

Cinque, G. (1993). A null theory of phrase and compound stress. *Linguistic Inquiry* 24: 239–267.

Crain, S. & R. Thornton. (1998). *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics.* Cambridge, Mass. MIT Press.

Cutler, A & J.A. Fodor (1979). Semantic focus and sentence comprehension. *Cognition* 7, 49-59.

Degen, J., P.A. Reeder, K. Carbary and M.K. Tanenhaus (2009). Using a Novel Experimental Paradigm to Investigate the Processing of Scalar Implicatures. Paper presented at Experimental Pragmatics 2009, Lyon, April 23-25.

De Neys, W., and W. Schaeken (2007) When People are More Logical Under Cognitive Load: Dual Task Impact on Scalar Implicature. *Experimental Psychology*, 54, 128–133.

Erickson, T. D., & M.E. Mattson (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, 20, 540–551.

Feeney, A., S. Scrafton, A. Duckworth and S.J. Handley (2004). The story of some: Everyday pragmatic inference by children and adults. *Canadian Journal of Experimental Psychology*, 58:90-1.

Ferreira, F., K. G. D. Bailey & V. Ferraro (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science* 11, 11–15.

Fox, D. 2007. Free choice and the theory of scalar implicatures. In: U. Sauerland and P. Stateva (eds.) *Presupposition and implicature in compositional semantics*: 71–120. New York: Palgrave Macmillan.

Fox, D. & B. Spector (2008). Economy and embedded exhaustification. Handout, available at http://lumiere.ens.fr/~bspector/Webpage/Fox_Spector_EE.pdf.

Frazier, L. (2009). Computing scalar implicatures. In: T. Friedman & S. Ito (eds.) *Proceedings of SALT 18*: 319-340. Ithaca, NY: CLC Publications.

Frazier, L. & C. Clifton (1998). Comprehension of Sluiced Sentences. *Language and Cognitive Processes* 13 (4), 499–520.

Fretheim, T. (1992) The Effect of Intonation on a Type of Scalar Implicature. *Journal of Pragmatics* 18, 1-30.

Gazdar G. (1979). *Pragmatics: Implicature, Presupposition, and Logical Form*. New York: Academic Press.

Geurts, B. (2006). Exclusive disjunction without implicatures. Ms. University of Nijmegen.

Geurts, B. (2009) & N. Pouscoulous. Embedded implicatures?!? *Semantics & Pragmatics* 2: 1–34.

Geurts, B. (in prep.) *Quantity Implicatures*. To appear with Cambridge University Press.

Grice, H. P. (1967/1975). Logic and conversation (from the William James lectures, Harvard University, 1967). In P. Cole & J. Morgan (eds.), *Syntax and semantics 3: Speech acts* (pp. 41-58). New York: Academic Press.

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

Grodner, D., N. Klein, K. Carbary and M. K. Tanenhaus (2007) Experimental evidence for rapid interpretation of pragmatic 'some'. Paper presented at XPRAG 2007, Berlin.

Groenendijk, Jeroen and Martin Stokhof (1984). *Studies on the Semantics of Questions and the Pragmatics of Answers*. Ph.D. Dissertation, University of Amsterdam.

Halliday, M. (1967). *Intonation and Grammar in British English*. Mouton, The Hague.

Hamblin, C.L. (1973). Questions in Montague English. *Foundations of Language* 10, 41-53.

Hausser, R. R. (1983) Surface Compositionality and the Semantics of Mood. In J. R. Searle, F. Kiefer, and M. Bierwisch (eds.), *Speech Act Theory and Pragmatics*, Reidel, Dordrecht.

Hirschberg, J. (1985) *A theory of Scalar Implicature*, Ph.D. thesis, UPenn.

de Hoop, H., & M. Kas. (1989). Sommige betekenisaspecten van enkele kwantoren, oftewel: enkele betekenisaspecten van sommige kwantoren. *TTT* 9:1. 31-49.

Horn, L. (1972). *On the Semantic Properties of Logical Operators in English*. Ph.D. diss., UCLA distributed by IULC.

Horn, L. (1989). *A Natural History of Negation*. Chicago: University of Chicago Press.

Horn, L. (2006) The Border Wars: A Neo-Gricean Perspective. In: K. von Heusinger and K.Turner (eds.) *Where Semantics Meets Pragmatics*, Elsevier, Oxford, England.

Huang, Y. & J. Snedeker, J. (2006). On-line interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Proceedings of the 28th annual meeting of the Cognitive Science Society*: 351-356.

Huang, Y. & J. Snedeker (2007). 'Logic & Conversation' revisited: Evidence for a division between semantic and pragmatic content in real time language comprehension. Ms. Harvard University.

Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, Mass. MIT Press.

Jacobs, J. (1983). *Fokus und Skalen. Zur Syntax und Semantik der Gradpartikel im Deutschen*. Tübingen: Niemeyer.

Katsos, N. (2006). *Experimental Investigations of the Effects of Structure and Context on the Generation of Scalar Implicatures*. PhD thesis. University of Cambridge. Cambridge.

Kiss, K.E. (1998). Identificational Focus versus Information Focus. *Language* 74.2: 245- 273.

Koornneef, A.W. (2008). *Eye catching Anaphora*. Ph.D. dissertation. Utrecht University: LOT Dissertation Series.

Lee, C. (2006) Contrastive (Predicate) Topic, Intonation, and Scalar Meanings. In: C. Lee, M. Gordon and D. Buring (eds.) *Topic and Focus: Crosslinguistic Perspectives on Meaning and Intonation*: 151-175. Springer.

van Kuppevelt, J. (1991). *Topic en Comment: Expliciete en Impliciete Vraagstelling in Discourse*. Ph.D. Thesis, University of Nijmegen.

van Kuppevelt, J. (1996). Inferring from topics. Scalar implicatures as topic-dependent inferences. *Linguistics and Philosophy* 19, 393-443.

Kratzer, A. (2003) Scalar Implicatures: Are There Any? Paper presented at the workshop "Polarity, Scalar Phenomena, and Implicatures", University of Milan – Bicocca, June 2003.

Kratzer, A. (2009). Situations in natural language semantics. *Stanford Encyclopedia of Philosophy*. Available at: http://plato.stanford.edu/entries/situations-semantics/.

Krifka, M. (1991) A Compositional Semantics for Multiple Focus Constructions. In: *Proceedings of SALT I*: 17–53.

Krifka, M. (1995) The Semantics and Pragmatics of Polarity Items. *Linguistic Analysis* 25(3-4), 209-257.

Krifka, M. (2002). Be brief and vague! And how bidirectional optimality theory allows for verbosity and precision. In D. Restle & D. Zaefferer (eds.), *Sounds and systems. Studies in structure and change. A festschrift for Theo Vennemann*: 439-458. Mouton de Gruyter Berlin.

Krifka, M. (2006) Association with focus phrases. In: V. Molnár & S. Winkler (eds.) *The architecture of focus*: 105-136. Berlin, Mouton de Gruyter.

Levinson, S.C. (1983) *Pragmatics*, Cambridge: Cambridge University Press.

Levinson, S.C. (2000). *Presumptive Meanings : the theory of generalized conversational implicature*. MIT Press. Cambridge, MA.

Mol, L., R. Verbrugge, & P. Hendriks (2004). Learning to reason about other people's minds. In: *Proceedings of the Joint Symposium on Virtual Social Agents*, SSAISB 2005 Convention, AISB, University of Hertfordshire, Hatfield.

Morris, R.K., & J.R. Folk (1998). Focus as a contextual priming mechanism in reading. *Memory & Cognition* 26, 1313-1322.

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

Noveck, I., G. Chierchia, F. Chevaux, R. Guelminger & E. Sylvestre (2002) Linguistic-pragmatic factors in interpreting disjunctions. *Thinking & Reasoning*, 8 (4), 297–326.

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

Panizza, D., G. Chierchia, and C. Clifton (2009). On the role of entailment patterns and scalar implicatures in the processing of numerals. *Journal of Memory and Language* 61: 503–518.

Papafragou A. & J. Musolino (2003). Scalar implicatures at the semantics-pragmatics interface. *Cognition* 80, 253-282.

Papafragou, A. & N. Schwarz (2006) *Most* Wanted. *Language Acquisition* 13(3), 207–251.

Pierrehumbert, J. & J. Hirschberg (1990). The meaning of intonational contours in the interpretation of discourse. In: P. Cohen, J. Morgan and M.E. Pollack (eds) *Intentions in Communication* 271–311.MIT Press Cambridge.

Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. In: *Papers in Semantics, OSU Working Papers in Linguistics*, Vol 49. Department of Linguistics, the Ohio State University, Columbus.
Van Rooij (2002)

van Rooij, R. & K. Schulz (2004). Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information* 13(4): 491-519.

Rooth, M. (1985). *Association with Focus*. PhD thesis, University of Massachusetts, GSLA, Amherst, MA.

Rooth, M. (1992). A theory of Focus Interpretation. *Natural Language Semantics* 1, 75-116.

Rooth, M. (1996). Focus. in S. Lappin (ed.) *The Handbook of Contemporary Semantic Theory*, Blackwell, London.

Sanford, A. J. & P. Sturt. (2002) Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences* 6:382–86.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27. 367-91.

Scharten, R. (1997). *Exhaustive interpretation: A discourse-semantic account*. Doctoral dissertation, Katholieke Universiteit, Nijmegen.

Schwarz, F., C. Clifton; and L. Frazier. (in progress). Strengthening 'or': Effects of focus and downward entailing contexts on scalar implicatures. UMass Amherst.

Schwarzschild, R. (1999). GIVENness, AVOIDF and Other Constraints on the Placement of Accents, *Natural Language Semantics* 7, 141-177.

Schulz, K. and R. van Rooij (2006) Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy* 29: 205-250.

Soames, S. (1982) How presuppositions are inherited: A solution to the projection problem. *Linguistic Inquiry* 13: 483–545.

Spector, B. (2007). Aspects of the Pragmatics of Plural Morphology. In: U. Sauerland and P. Stateva (eds.) *Presupposition and implicature in compositional semantics*: 243-281. Palgrave-Macmillan.

Sperber, D. & D. Wilson (1986). *Relevance: Communication and Cognition*, Harvard University Press, Cambridge, MA.

Sperber, D. & D. Wilson (1995) "Postface." In: D. Sperber and D. Wilson, *Relevance: Communication and Cognition*, second edition. Oxford: Blackwell.

Stalnaker, R. C. (1972). Pragmatics. In D. Davidson & G. Harman (eds.), *Semantics of Natural Language*: 380-397. Dordrecht: North-Holland.

Storto, G. & Tanenhaus, M. K. (2005) Are scalar implicatures computed online? *Proceedings of SuB 9*: 431-445.

Sturt, P., A.J. Sanford, A. Stewart & E.J. Dawydiak (2004). Linguistic focus and good-enough representations: an application of the change-detection paradigm. *Psychonomic Bulletin and Review* 11: 882-888.

von Stechow, A. (1990). Focusing and backgrounding operators. In W. Abraham (ed.) *Discourse Particles*, 6 (pp. 37–84). Amsterdam: John Benjamins.

de Swart, H. (1998) *Introduction to Natural Language Semantics*. Stanford: CSLI Publications.

Tanenhans, M.K. M.J. Spivey-Knowlton, K.M. Eberhard & J.C. Sedivy (1995). Integration of visual and linguistic information during spoken language comprehension. *Science* 268: 1632-1634.

Vallduví, E. (1990) *The information component*. Ph.D. Thesis, University of Pennsylvania.Veenker (2000)

Ward, P., P. Sturt, P. (2007). Linguistic focus and memory: an eye movement study. *Memory and Cognition* 35, 73–86.

Zondervan, A. (2007). Effects of question under discussion and focus on scalar implicatures. *Proceedings of the Fifth Semantics in the Netherlands Day*: 39-52.

Zondervan, A. (2009). Experiments on QUD and focus as a contextual constraint on scalar implicature calculation. In: U. Sauerland and K. Yatsushiro (eds.) *Semantics and Pragmatics, From Experiment to Theory,* Palgrave Macmillan.

Zondervan, A. (2010) The role of QUD and focus on the scalar implicature of *most*. In: Meibauer, J. and M. Steinbach, (eds.) *Experimental Semantics/Pragmatics*, Benjamins.

APPENDIX 1

Items Experiment 1 (chapter 3)

C1 = focus, C2 = non-focus

1. (C1) Harry ging naar een huisfeest van een vriendin van hem. Hij had beloofd nog wat versnaperingen mee te nemen voor het feest. Zij hoopte dat hij nootjes zou meenemen, want die was ze vergeten te kopen. Harry had stokbrood meegenomen. Hij had ook chips meegenomen. Hij had nog even overwogen nootjes mee te nemen, maar dat leek hem toch niet nodig.
A: "Wat had Harry meegenomen?"
B: "Harry had stokbrood of chips meegenomen."

1. (C1) Harry went to a party of a friend (fem.) of his. He had promised to bring some snacks for the party. She was hoping he would bring nuts, because she forgot to buy those. Harry brought (French) bread. He also brought potato chips. He had considered brining nuts for a while, but he figured it would not be necessary.
A: "What did Harry bring?"
B: "Harry brought bread or chips."

1. (C2) Harry en Hermelien gingen naar een huisfeest van Joost, een vriend van hen. Joost had hen allebei vantevoren gevraagd om stokbrood of chips mee te nemen. Hij had gezegd dat hij degene die stokbrood of chips mee zou nemen, eeuwig dankbaar zou zijn. Harry had stokbrood meegenomen. Hij had ook chips meegenomen. Hermelien was normaal gesproken zeer betrouwbaar, maar dit keer was ze straal vergeten ook maar iets mee te nemen.
A: "Wie had er stokbrood of chips meegenomen?"
B: "Harry had stokbrood of chips meegenomen."

1. (C2) Harry and Hermione went to a party of Joost, a friend of theirs. Joost had asked them both beforehand to bring (French) bread or potato chips. He had told them that he would be forever grateful to the one who would bring bread or chips. Harry brought bread. He also brought chips. Hermione was normally very reliable, but this time she completely forgot to bring anything at all.
A: "Who brought bread or chips?"
B: "Harry brought bread or chips."

2. (C1) Adriaan was aan het werk in de meubelmakerij. Hij had zijn vrouw beloofd vandaag wat mooie dingen te maken, maar het was een verrassing wat hij zou maken. Zijn vrouw hoopte dat hij een bank zou maken, want die vond ze altijd erg mooi. Adriaan maakte om te beginnen een kast. Even later maakte hij ook nog een stoel. Hij wilde ook nog een bank maken, maar dat werd hem te veel werk.
A: "Wat heeft Adriaan gemaakt?"
B: "Adriaan heeft een kast of een stoel gemaakt."

2. (C1) Adriaan was working in the furniture-making shop. He had promised his wife to make some beautiful things today, but what he would make would be a surprise. His wife was hoping he would make a couch, because she always liked his couches. First Adriaan made a cupboard. A while later he also made a chair. He wanted to make a couch too, but in the end he decided that would be too much work.
A: "What did Adriaan make?"
B: "Adriaan made a cupboard or a chair."

2. (C2) Bassie en Adriaan waren aan het werk in de meubelmakerij. Hun baas had ze allebei de opdracht gegeven vandaag een kast of een stoel te maken. Hij had gezegd dat degene die vandaag een kast of een stoel zou maken, morgen vrij zou zijn. Adriaan had al gauw een kast af. Even later had hij ook een stoel af. Bassie was zo onhandig dat het hem niet lukte ook maar iets in elkaar te zetten.
A: "Wie heeft er een kast of een stoel gemaakt?"
B: "Adriaan heeft een kast of een stoel gemaakt."

2. (C2) Bassie and Adriaan were working in the furniture-making shop. Their boss had given them both the assignment to make a cupboard or a chair today. He had told them that the one who would make a cupboard or a chair today, would get the day off tomorrow. Adriaan soon finished a cupboard. A while later he also finished a chair. Bassie was so clumsy that he didn't manage to put anything together.
A: "Who made a cupboard or a chair?"
B: "Adriaan made a cupboard or a chair."

3. (C1) Gerard was aan het jagen voor het kerstdiner bij zijn moeder. Hij had zijn moeder beloofd wat smakelijk wild te schieten in het bos. Zijn moeder had gezegd dat als hij een ree zou schieten, hij niet zou hoeven afwassen. Gerard schoot eerst een eend. Even later schoot hij een konijn. Hij zag nergens een ree dus die kon hij ook niet schieten.
A: "Wat heeft Gerard geschoten?"
B: "Gerard heeft een eend of een konijn geschoten."

3. (C1) Gerard was hunting for game for the Christmas dinner at his mother's. He had promised his mother to shoot some tasty game in the forest. His mother had said that if he would shoot a deer, he wouldn't have to do the dishes. Gerard shot a duck first. A while later he also shot a rabbit. He didn't see a deer anywhere so he couldn't shoot one.
A: "What did Gerard shoot?"
B: "Gerard shot a duck or a rabbit."

3. (C2) Gerard en Gordon waren aan het jagen bij het landhuis van Gordons moeder. Zij had ze allebei de opdracht gegeven een eend of een konijn te schieten. Ze had beloofd dat degene die een eend of een konijn zou schieten, niet zou hoeven af te wassen 's avonds. Gerard had geluk en schoot direct een eend. Even later schoot hij ook een konijn. Gordon had geen geluk en raakte helemaal niks.
A: "Wie heeft er een eend of een konijn geschoten?"
B: "Gerard heeft een eend of een konijn geschoten."

3. (C2) Gerard and Gordon were hunting near the mansion of Gordon's mother. She had given both of them the assignment to shoot a duck or a rabbit. She had promised that the one who would shoot a duck or a rabbit, wouldn't have to do the dishes that night. Gerard was lucky and he shot a duck right away. A little later he also shot a rabbit. Gordon was unlucky and he didn't shoot anything at all.
A: "Who shot a duck or a rabbit?"
B: "Gerard shot a duck or a rabbit."

4. (C1) Katja was zeedieren aan het zoeken op het strand bij haar grootouders. Ze had haar opa beloofd wat mooie dieren te vinden. Hij had gezegd dat als ze een oester zou vinden, ze een tientje zou krijgen. Katja vond al snel een krab. Even later vond ze ook een zeester. Maar hoe hard ze ook zocht, ze vond geen oester.
A: "Wat heeft Katja gevonden?"
B: "Katja heeft een krab of een zeester gevonden."

4. (C1) Katja was searching for marine animals on the beach near her grandparents' house. She had promised her grandfather to find some beautiful animals. He had said that if she would find an oyster, she would get ten bucks. Katja soon found a crab. Not much later she also found a starfish. But no matter how hard she looked, she didn't find an oyster.
A: "What did Katja find?"
B: "Katja found a crab or a starfish."

4. (C2) Katja en Birgit gaan zeedieren zoeken op het strand bij hun grootouders. Hun opa had ze aangemoedigd allebei op zoek te gaan naar een krab of een zeester. Hij had ze beloofd dat degene die een krab of een zeester zou vinden, een tientje zou krijgen. Na enig zoeken vond Katja een krab. Niet veel later vond ze ook een zeester. Birgit kon niks vinden en moest met lege handen naar huis terugkeren.
A: "Wie heeft er een krab of een zeester gevonden?"
B: "Katja heeft een krab of een zeester gevonden."

4. (C2) Katja and Birgit were searching for marine animals on the beach near their grandparents' house. Their grandfather had encouraged them both to go look for a crab or a starfish. He had promised them that the one who would find a crab or a starfish, would get ten bucks. After some searching Katja found a crab. Not much later she also found a starfish. Birgit couldn't find anything and had to return to the house empty handed.
A: "Who found a crab or a starfish?"
B: "Katja found a crab or a starfish."

5. (C1) Barrie ging naar de Jamin. Hij had zijn klasgenoot Sjonnie beloofd wat snoep voor hem te jatten. Sjonnie hoopte dat hij kauwgomballen zou jatten, zodat hij stoer bellen zou kunnen blazen. In de Jamin jatte Barrie een chocoladereep. Hij jatte ook een zak drop. Hij wilde eigenlijk ook nog een zak zuurtjes jatten, maar dat duurde hem te lang.
A: "Wat heeft Barrie gejat?"
B: "Barrie heeft een zak drop of een chocoladereep gejat."

5. (C1) Barrie went to the candy store. He had promised his classmate Sjonnie to steal some candy for him. Sjonnie was hoping he would steal gumballs, so he could blow bubbles. In the candy store Barrie stole a chocolate bar. He also stole a bag of liquorice. He actually also wanted to steal a bag of sourballs, but he decided it was taking too long.
A: "What did Barrie steal?"
B: "Barrie stole a bag of liquorice or a chocolate bar."

5. (C2) Barrie en Mitchel gingen naar de Jamin. Hun klasgenoot Sjonnie had ze uitgedaagd een chocoladereep of een zak drop voor hem te jatten. Hij had gezegd dat degene die voor hem een chocoladereep of een zak drop zou jatten, een dag zijn brommer zou mogen lenen. In de Jamin jatte Barrie een chocoladereep. Hij stal ook een zak drop. Mitchel durfde uiteindelijk toch niet en liep gauw de winkel uit.
A: "Wie heeft er een zak drop of een chocoladereep gejat?"
B: "Barrie heeft een zak drop of een chocoladereep gejat."

5. (C2) Barrie and Mitchel went to the candy store. Their classmate Sjonnie had challenged them to steal a chocolate bar or a bag of liquorice for him. He had told them that the person who would steal a chocolate bar or a bag of liquorice for him, would get to borrow his scooter for a day. In the candy store Barrie stole a chocolate bar. He also stole a bag of liquorice. Mitchel was eventually too afraid and quickly walked out of the store.
A: "Who stole a bag of liquorice or a chocolate bar?"
B: "Barrie stole a bag of liquorice or a chocolate bar."

6. (C1) Jantje ging op schoolreisje. Hij had van zijn moeder een hele zak fruit meegekregen. Ze hoopte vooral dat hij een kiwi op zou eten, want dat is volgens haar de gezondste vrucht. Op schoolreisje at Jantje op de heenweg een appel. Op de terugweg at hij een peer. Later wilde hij ook nog een kiwi eten, maar hij was zijn zak fruit kwijtgeraakt.
A: "Wat heeft Jantje gegeten?"
B: "Jantje heeft een appel of een peer gegeten."

6. (C1) Jantje went on a school trip. His mother had given him a bag full of fruit. She was especially hoping he would eat a kiwi, because she thought it was the most healthy type of fruit. At the school trip Jantje ate an apple on the way out. On the way back he ate a pear. Later on he also wanted to eat a kiwi, but he had lost his bag of fruit.
A: "What did Jantje eat?"
B: "Jantje ate an apple or a pear."

6. (C2) De broertjes Jantje en Pietje wilden allebei vandaag laat opblijven om hun favoriete tv-programma te zien. Hun moeder wilde dat ze meer fruit zouden eten, dus ze had gezegd dat degene die een appel of een peer zou eten, laat op zou mogen blijven. Jantje at direct een appel op. Even later at hij ook een peer. Pietje haatte fruit en besloot niks te eten.
A: "Wie heeft een appel of een peer gegeten?"
B: "Jantje heeft een appel of een peer gegeten."

6. (C2) The brothers Jantje and Pietje both wanted to stay up late today to watch their favorite TV-show. Their mother wanted them to eat more fruit, so she told them that the one who would eat an apple or a pear, would get to stay up late. Jantje immediately ate an apple. A little later he also ate a pear. Pietje hated fruit and decided not to eat anything.
A: "Who ate an apple or a pear?"
B: "Jantje ate an apple or a pear."

Exp 1 Average SI-rates per item

item	C1 [focus]	C2 [non-focus]
1	88%	53%
2	68%	59%
3	68%	59%
4	53%	37%
5	79%	65%
6	82%	58%

APPENDIX 2

Items Experiment 2 (chapter 3)
C1 = focus, C2 = non-focus

1. (C1) Julie en Karin waren samen zeedieren aan het zoeken op het strand. Na enig zoeken vond Julie een krab. Niet veel later vond ze ook een zeester. Karin vond helaas helemaal niks. Toen Karin terugkwam, vroeg haar moeder wat voor zeedieren Julie had gevonden. Karin antwoordde dat Julie een krab of een zeester had gevonden.

1. (C1) Julie and Karin were searching for marine animals on the beach. After some searching Julie found a crab. Not much later she also found a starfish. Unfortunately, Karin didn't find anything. When Karin returned, her mother asked what kind of marine animals Julie had found. Karin answered that Julie had found a crab or a starfish.

1. (C2) Julie en Karin waren samen zeedieren aan het zoeken op het strand. Na enig zoeken vond Julie een krab. Niet veel later vond ze ook een zeester. Karin vond helaas helemaal niks. Toen ze terugkwamen, vroeg hun moeder wie van hen een krab of een zeester had gevonden. Karin antwoordde dat Julie een krab of een zeester had gevonden.

1. (C2) Julie and Karin were searching for marine animals on the beach. After some searching Julie found a crab. Not much later she also found a starfish. Unfortunately, Karin didn't find anything. When they returned, their mother asked who had found a crab or a starfish. Karin answered that Julie had found a crab or a starfish.

2. (C1) Jantje en Pietje gingen op schoolreisje. Ze hadden allebei van de leraar een zak fruit gekregen. Jantje at een appel. Even later at hij ook een kiwi. Pietje vond fruit maar vies en hij at niks. Toen Pietje terugkwam, vroeg de leraar wat voor fruit Jantje had gegeten. Pietje zei dat Jantje een appel of een kiwi had gegeten.

2. (C1) Jantje and Pietje went on a school trip. The teacher had given them both a bag of fruit. Jantje ate an apple. A little later he also ate a kiwi. Pietje didn't like fruit at all and he didn't eat anything. When Pietje returned, the teacher asked what kind of fruit Jantje had eaten. Pietje answered that Jantje had eaten an apple or a kiwi.

2. (C2) Jantje en Pietje gingen op schoolreisje. Ze hadden allebei van de leraar een zak fruit gekregen. Jantje at een appel. Even later at hij ook een kiwi. Pietje vond fruit maar vies en hij at niks. Toen ze terugkwamen, vroeg de leraar wie van hen een appel of een kiwi had gegeten. Pietje zei dat Jantje een appel of een kiwi had gegeten.

2. (C2) Jantje and Pietje went on a school trip. The teacher had given them both a bag of fruit. Jantje ate an apple. A little later he also ate a kiwi. Pietje didn't like fruit at all and he didn't eat anything. When they returned, the teacher asked who of them had eaten an apple or a kiwi. Pietje answered that Jantje had eaten an apple or a kiwi.

3. (C1) De hooligans Sjaak en Willem waren in de stad vernielingen aan het aanrichten. Sjaak vernielde een lantaarnpaal. Even later vernielde hij ook een verkeersbord. Willem durfde uiteindelijk toch niks te vernielen. Toen Willem daarna in het supportershome kwam, vroeg een mede-hooligan wat voor dingen Sjaak had vernield. Willem antwoordde dat Sjaak een lantaarnpaal of een verkeersbord had vernield.

3. (C1) The hooligans Sjaak and Willem were vandalizing things in the city. Sjaak vandalized a lamppost. A little later he also vandalized a traffic sign. Willem eventually was too afraid to vandalize anything. When Willem got to the supporters meeting place afterwards, a fellow hooligan asked what things Sjaak had vandalized. Willem answered that Sjaak had vandalized a lamppost or a traffic sign.

3. (C2) De hooligans Sjaak en Willem waren in de stad vernielingen aan het aanrichten. Sjaak vernielde een lantaarnpaal. Even later vernielde hij ook een verkeersbord. Willem durfde uiteindelijk toch niks te vernielen. Toen ze daarna in het supportershome kwamen, vroeg een mede-hooligan wie van hen een lantaarnpaal of een verkeersbord had vernield. Willem antwoordde dat Sjaak een lantaarnpaal of een verkeersbord had vernield.

3. (C2) The hooligans Sjaak and Willem were vandalizing things in the city. Sjaak vandalized a lamppost. A little later he also vandalized a traffic sign. Willem eventually was too afraid to vandalize anything. When they got to the supporters meeting place afterwards, a fellow hooligan asked who of them had vandalized a lamppost or a traffic sign. Willem answered that Sjaak had vandalized a lamppost or a traffic sign.

4. (C1) Jessica Simpson en Nick Lachey gingen voor de verandering zelf het huis schoonmaken. Nick maakte het toilet schoon. Even later maakte hij ook de douche schoon. Jessica leek het toch allemaal teveel werk en ze maakte uiteindelijk niks schoon. Toen Jessica daarna een vriendin over de schoonmaakactie vertelde, vroeg die wat Nick had schoongemaakt. Jessica antwoordde dat Nick de douche of het toilet had schoongemaakt.

4. (C1) Jessica Simpson and Nick Lachey decided to clean the house themselves for a change. Nick cleaned the toilet. A little later he also cleaned the shower. Jessica eventually decided it was too much work and she didn't clean anything. When Jessica told a friend (fem.) about the cleaning activities, she asked what Nick had cleaned. Jessica answered that Nick had cleaned the shower or the toilet.

4. (C2) Jessica Simpson en Nick Lachey gingen voor de verandering zelf het huis schoonmaken. Nick maakte het toilet schoon. Even later maakte hij ook de douche schoon. Jessica leek het toch allemaal teveel werk en ze maakte uiteindelijk niks schoon. Toen ze daarna naar een vriend over de schoonmaakactie vertelden, vroeg die wie van hen de douche of het toilet had schoongemaakt. Jessica antwoordde dat Nick de douche of het toilet had schoongemaakt.

4. (C2) Jessica Simpson and Nick Lachey decided to clean the house themselves for a change. Nick cleaned the toilet. A little later he also cleaned the shower. Jessica eventually decided it was too much work and she didn't clean anything. When they told a friend about the cleaning activities, he asked who of them had cleaned the shower or the toilet. Jessica answered that Nick had cleaned the shower or the toilet.

5. (C1) André van Duin en Ron Brandsteder waren aan het golfen. Ron maakte een birdie. Een paar holes later maakte hij zelfs een eagle. André had zijn dag niet en maakte alleen maar bogeys. Toen André in het clubhuis kwam, vroeg de eigenaar wat voor scores Ron had gemaakt. André antwoordde dat Ron een birdie of een eagle had gemaakt.

5. (C1) André van Duin and Ron Brandsterer were playing golf. Ron made a birdie. A couple of holes later he even made an eagle. André had an off-day and only made bogeys. When André entered the club house, the owner asked what kind of scores Ron had made. André answered that Ron had made a birdie or an eagle.

5. (C2) André van Duin en Ron Brandsteder waren aan het golfen. Ron maakte een birdie. Een paar holes later maakte hij zelfs een eagle. André had zijn dag niet en maakte alleen maar bogeys. Toen ze in het clubhuis kwamen, vroeg de eigenaar wie van hen een birdie of een eagle had gemaakt. André antwoordde dat Ron een birdie of een eagle had gemaakt.

5. (C2) André van Duin and Ron Brandsterer were playing golf. Ron made a birdie. A couple of holes later he even made an eagle. André had an off-day and only made bogeys. When they entered the club house, the owner asked who of them had made a birdie or an eagle. André answered that Ron had made a birdie or an eagle.

6. (C1) Anton en Henk waren aan het eten in een Japans restaurant. Anton nam sushi. Even later nam hij ook wat sashimi. Henk had geen zin in vis en nam teriyaki met kip. Toen Henk thuiskwam, vroeg zijn huisgenoot wat Anton had genomen. Henk antwoordde dat Anton sushi of sashimi had genomen.

6. (C1) Anton and Henk were having dinner at a Japanese restaurant. Anton took sushi. A while later he also took some sashimi. Henk did not feel like having fish and took chicken teriyaki. When Henk returned home, his housemate asked what Anton had taken. Henk answered that Anton took sushi or sashimi.

6. (C2) Anton en Henk waren aan het eten in een Japans restaurant. Anton nam sushi. Even later nam hij ook wat sashimi. Henk had geen zin in vis en nam teriyaki

met kip. Toen ze thuiskwamen, vroeg hun huisgenoot wie van hen sushi of sashimi had genomen. Henk antwoordde dat Anton sushi of sashimi had genomen.

6. (C2) Anton and Henk were having dinner at a Japanese restaurant. Anton took sushi. A while later he also took some sashimi. Henk did not feel like having fish and took chicken teriyaki. When they returned home, their housemate asked who of them had taken sushi or sashimi. Henk answered that Anton took sushi or sashimi.

Exp 2 Average SI-rates per item

item	C1 [focus]	C2 [non-focus]
1	61%	35%
2	70%	48%
3	65%	52%
4	70%	39%
5	61%	39%
6	74%	30%

APPENDIX 3

Items Experiment 3 (chapter 3)

C1 = focus, C2 = non-focus

1. (C1) Paola ging lunch halen in de kantine. In het fruitvak lagen bananen, sinaasappels, appels en peren.
Bananen vond Paola niet zo lekker, dus ze nam geen banaan.
De sinaasappels zagen er een beetje beurs uit, dus die vielen ook af.
De appels zagen er sappig uit, dus Paola nam er één.
Een peer had Paola al in geen jaren gehad, dus ze besloot voor de verandering ook een peer te nemen.

"Paola heeft EEN APPEL OF EEN PEER uit het fruitvak genomen."

1. (C1) Paola was getting lunch at the cafeteria. At the fruit section there were bananas, oranges, apples and pears.
Bananas Paola didn't like, so she didn't take any.
The oranges looked a bit old, so she also didn't take any of those.
The apples looked nice and juicy, so Paola took one.
A pear Paola hadn't eaten in years, so for a change she decided to also take a pear.

"Paola took AN APPLE OR A PEAR from the fruit section."

1. (C2) Paola, Linda, Betty en Ginger waren lunch aan het halen in de kantine. In het fruitvak lagen appels en peren.
Paola was een echte gezondheidsfreak dus ze nam een appel. Na enig nadenken nam ze ook een peer.
Linda had al zoveel andere dingen gepakt, dat ze besloot geen fruit te nemen.
Betty hield überhaupt niet van fruit, dus ze liet de appels en de peren links liggen.
Ginger overwoog even een peer te nemen, maar ze wist niet zeker of ze genoeg geld bij zich had, dus deed ze het niet.

"PAOLA heeft een appel of een peer uit het fruitvak genomen."

1. (C2) Paola, Linda, Betty and Ginger were getting lunch at the cafeteria. At the fruit section there were only apples and pears.
Paola was a real health freak, so she took an apple. After some consideration she also took a pear.
Linda already got a lot of other food, so she decided not to take any fruit.
Betty didn't like fruit at all, so she ignored the apples and the pears.
Ginger considered taking a pear, but she wasn't sure whether she brought enough money, so she didn't take any.

"PAOLA took an apple or a pear from the fruit section."

2. (C1) George ging naar een verkiezingsbijeenkomst. Bij de ingang lagen buttons, vlaggetjes, bordjes en petjes met de naam van de kandidaat.
Een button vond George wel gaaf, dus die pakte hij.
De vlaggetjes vond George geweldig, dus hij pakte er snel één.
Een bordje leek George maar onhandig groot, dus die nam hij niet.
De petjes vond George een beetje kinderachtig, dus die nam hij ook niet.

"George heeft EEN VLAGGETJE OF EEN BUTTON gepakt bij de ingang."

2. (C1) George went to an election rally. At the entrance there were buttons, little flags, signs and caps with the name of the candidate on them.
A button George thought was pretty cool, so he took it.
The flags George thought were amazing, so he quickly took one.
A sign George thought was inconveniently big, so he didn't take it.
The caps George thought were a bit childish, so he also didn't take one.

"George took A FLAG OR A BUTTON at the entrance."

2. (C2) John, George, Bill en Ronald gingen naar een verkiezingsbijeenkomst. Bij de ingang lagen buttons en vlaggetjes met de naam van de kandidaat.
John vond buttons en vlaggetjes altijd een beetje kinderachtig, dus hij pakte niks.
George wilde de kandidaat graag steunen, dus hij pakte meteen een vlaggetje. Hij pakte ook een button.
Bill hield best van buttons en vlaggetjes, maar deze vond hij echt te lelijk, dus hij pakte niks.
Ronald vond buttons en vlaggetjes sowieso onzin, dus hij pakte niks.

"GEORGE heeft een vlaggetje of een button gepakt bij de ingang."

2. (C2) John, George, Bill and Ronald went to an election rally. At the entrance there were buttons and little flags with the name of the candidate on them.
John thought buttons and flags were a bit childish, so he didn't take anything.
George wanted to support the candidate, so he immediately took a flag. He also took a button.
Bill liked buttons and flags in general, but these were just too ugly, so he didn't take anything.
Ronald thought buttons and flags were nonsense anyway, so he didn't take anything.

"GEORGE took a flag or a button at the entrance."

3. (C1) Dieter was op een beurs. Bij een stand werden gratis pennen, blocnotes, sleutelhangers en flessenopeners weggegeven.
Pennen zijn altijd handig, dus die nam Dieter aan.
Blocnotes had Dieter al genoeg, dus die nam hij niet aan.
De sleutelhanger vond Dieter er wel aardig uitzien, dus die nam hij wel.
De flessenopener zag er onbetrouwbaar uit, dus die nam hij niet.

"Dieter heeft EEN PEN OF EEN SLEUTELHANGER aangenomen bij de stand."

3. (C1) Dieter was at a convention. At one stand people were giving out pens, writing pads, keychains and bottle openers.
Pens always come in handy, so Dieter took it.
Writing pads Dieter already had loads of, so he didn't take it.
The keychain Dieter thought looked okay, so he did take that.
The bottle opener didn't look very sturdy, so he didn't take it.

"Dieter took A PEN OR A KEYCHAIN at the stand."

3. (C2) Helmut, Heinz, Franz en Dieter waren op een beurs. Bij een stand werden gratis pennen en sleutelhangers weggegeven.
Helmut hoefde die gratis troep niet, dus nam hij niks aan.
Heinz vond de pennen en de sleutelhangers maar lelijk, dus hij nam ook niks aan.
Franz vond die plastic dingen milieuonvriendelijk, dus hij nam niks aan.
Dieter was gek op gratis dingen, dus hij nam een pen aan. Toen hij ook een sleutelhanger aangeboden kreeg, nam hij die ook aan.

"DIETER heeft een pen of een sleutelhanger aangenomen bij de stand."

3. (C2) Helmut, Heinz, Franz and Dieter were at a convention. At one stand they were giving out pens and keychains.
Helmut didn't want that free junk, he he didn't take anything.
Heinz thought the pens and the keychains were ugly, so he didn't take anything either.
Franz felt those plastic things were bad for the environment, so he didn't take any.
Dieter loved free stuff, so he took a pen. When they also offered him a keychain, he also took that.

"DIETER took a pen or a keychain at the stand."

Exp 3 Average SI-rates per item

item	Version 1 [focus]	Version 2 [non-focus]
1	90%	56%
2	85%	56%
3	80%	52%

APPENDIX 4

Items of Experiment 4 (chapters 4 and 8) and Experiment 6 (chapters 4 and 9)

C1 = [−alternative relevance, +focus]
C2 = [+alternative relevance, +focus]
C3 = [−alternative relevance, −focus]
C4 = control

Items of conditions 1 and 2 are represented together, with the difference in boldface (not in actual items). The items of condition 4 are not given separately, as they are identical to condition 3, except for having a term answer instead of a sentential answer.

Reported are the items and the total question-set per item, from which the questions were used as follows: All items in Experiment 4 contained three questions in the following order: The comprehension question (CQ) (which varied over conditions), the speaker expertise question (IgnI-Q), and the Scalar Implicature question (SI-Q). (These labels were absent in the actual items)

In Experiment 6 all items contained two questions. As this was a between-subjects design, there were two versions of the experiment. In the first version all items contained the comprehension question (CQ) and the SI-question (SI-Q) (in that order). In the exhaustivity version all items contained the comprehension question and the external exhaustivity question (Exh-Q).

Items start on the next page. I only provide complete translations for items 1-6. For items 7-12 I only translate condition 1/condition 2, as from the translations of items 1-6 it should be obvious how conditions 3 and 4 and the questions are formed from the condition 1/condition 2 story.

1. (C1/C2) Marieke vertelde aan haar moeder dat Laura gisteren zeedieren was gaan zoeken op het strand, en dat haar vader had gezegd dat als ze **een krab of een mossel** (C1) / **minstens twee zeedieren** (C2) zou vinden, ze die avond later op zou mogen blijven. De moeder zei: "Oh, en wat heeft Laura gevonden?" Marieke antwoordde: "Ze heeft een krab of een mossel gevonden."

*1. (C1/C2) Marieke told her mother that Laura went searching for marine animals on the beach yesterday, and that her father had told her that if she would find **a crab or a mussel** (C1) / **at least two marine animals** (C2), she would get to stay up late that night. The mother said: "Oh, and what did Laura find?" Marieke answered: "She found a crab or a mussel."*

1. (C3) Marieke vertelde aan haar moeder dat Laura en Barbara gisteren zeedieren waren gaan zoeken op het strand, en dat hun vader had gezegd dat degene die een krab of een mossel zou vinden, die avond later op zou mogen blijven. De moeder zei: "Oh, en wie van hen heeft een krab of een mossel gevonden?" Marieke antwoordde: "Laura heeft een krab of een mossel gevonden."

1. (C3) Marieke told her mother that Laura and Barbara went searching for marine animals on the beach yesterday, and that their father had told them that the one who would find a crab or a mussel, would get to stay up late that night. The mother said: "Oh, and who of them found a crab or a mussel?" Marieke answered: "Laura found a crab or a mussel."

<u>Question-set item 1</u>

CQ (C1):	Wat zou Laura mogen als ze een krab of een mossel zou vinden?
CQ (C2):	Wat zou Laura mogen als ze minstens twee zeedieren zou vinden?
CQ (C3+C4):	Wat zouden Laura en Barbara mogen als ze een krab of een mossel zou vinden?
IgnI-Q:	Denk je dat Marieke precies weet wat Laura heeft gevonden? *Do you think Marieke knows exactly what Laura found?*
SI-Q:	Denk je dat het mogelijk is dat Laura zowel een krab als een mossel heeft gevonden? *Do you think it is possible that Laura found both a crab and a mussel?*
Exh-Q:	Denk je dat het mogelijk is dat Laura ook nog iets anders dan een krab of een mossel heeft gevonden? *Do you think it is possible that Laura also found something else than a crab or a mussel?*

2. (C1/C2) Nathalie vertelde aan een vriendin dat Chris gisteren naar de snackbar was geweest, maar dat zijn vader van tevoren had gezegd dat als hij **een kroket of een frikandel** (C1) / **meer dan één snack** (C2) zou nemen, hij geen avondeten zou krijgen. De vriendin zei: "Oh, en wat heeft Chris genomen?" Nathalie antwoordde: "Hij heeft een kroket of een frikandel genomen."

2. (C1/C2) Nathalie told a friend (fem.) that Chris went to the snack bar yesterday, but that his father had said beforehand that if he would take ***a croquette or a meat roll (C1) / more than one snack (C2)****, he would not get dinner. The friend said: "Oh, and what did Chris take?" Nathalie answered: "He took a croquette or a meat roll."*

2. (C3) Nathalie vertelde aan een vriendin dat Chris en Marco gisteren naar de snackbar waren geweest, maar dat hun vader van tevoren had gezegd dat degene die een kroket of een frikandel zou nemen, geen avondeten zou krijgen. De vriendin zei: "Oh, en wie van hen heeft een kroket of een frikandel genomen?" Nathalie antwoordde: "Chris heeft een kroket of een frikandel genomen."

2. (C3) Nathalie told a friend (fem.) that Chris and Marco went to the snack bar yesterday, but that their father had said beforehand that the one who would take a croquette or a meat roll would not get dinner. The friend said: "Oh, and who of them took a croquette or a meat roll?" Nathalie answered: "Chris took a croquette or a meat roll."

Question-set item 2

CQ (C1):	Wie had gezegd dat als Chris een kroket of een frikandel zou nemen, hij geen avondeten zou krijgen?
CQ (C2):	Wie had gezegd dat als Chris meer dan één snack zou nemen, hij geen avondeten zou krijgen?
CQ (C3+C4):	Wie had gezegd dat degene die een kroket of een frikandel zou nemen, geen avondeten zou krijgen?
IgnI-Q:	Denk je dat Nathalie precies weet wat Chris heeft genomen? *Do you think Nathalie knows exactly what Chris took?*
SI-Q:	Denk je dat het mogelijk is dat Chris zowel een kroket als een frikandel heeft genomen? *Do you think it is possible that Chris took both a croquette and a meat roll?*
Exh-Q:	Denk je dat het mogelijk is dat Chris ook nog iets anders dan een kroket of een frikandel heeft genomen? *Do you think it is possible that Chris also took something else than a croquette or a meat roll?*

3. (C1/C2) Astrid vertelde aan haar moeder dat Patrick gisteren was gaan vissen bij het meer, en dat zijn vader van tevoren had gezegd dat als hij **een karper of een snoek (C1) / minstens twee vissen (C2)** zou vangen, hij een tientje zou krijgen. Astrids moeder zei: "Oh, en wat heeft Patrick gevangen?" Astrid antwoordde: "Hij heeft een karper of een snoek gevangen."

*3. (C1/C2) Astrid told her mother that Patrick went fishing at the lake yesterday, and that his father had said beforehand that if he would catch **a carp or a pike (C1) / at least two fish (C2)**, he would get ten bucks. Astrids mother said: "Oh, and what did Patrick catch?" Astrid answered: "He caught a carp or a pike."*

3. (C3) Astrid vertelde aan haar moeder dat Patrick en Maurice gisteren waren gaan vissen bij het meer, en dat hun vader van tevoren had gezegd dat degene die een karper of een snoek zou vangen, een tientje zou krijgen. Astrids moeder zei: "Oh, en wie van hen heeft een karper of een snoek gevangen?" Astrid antwoordde: "Patrick heeft een karper of een snoek gevangen."

3. (C3) Astrid told her mother that Patrick and Maurice went fishing at the lake yesterday, and that their father had said beforehand that the one who would catch a carp or a pike, would get ten bucks. Astrids mother said: "Oh, and who of them caught a carp or a pike?" Astrid answered: "Patrick caught a carp or a pike."

Question-set item 3

CQ (C1):	Wie had gezegd dat als Patrick een karper of een snoek zou vangen, hij een tientje zou krijgen?
CQ (C2):	Wie had gezegd dat als Patrick minstens twee vissen zou vangen, hij een tientje zou krijgen?
CQ (C3+C4):	Wie had gezegd dat degene die een karper of een snoek zou vangen, een tientje zou krijgen?
IgnI-Q:	Denk je dat Astrid precies weet wat Patrick heeft gevangen? *Do you think Astrid knows exactly what Patrick caught?*
SI-Q:	Denk je dat het mogelijk is dat Patrick zowel een karper als een snoek heeft gevangen? *Do you think it is possible that Patrick caught both a carp and a pike?*
Exh-Q:	Denk je dat het mogelijk is dat Patrick ook nog iets anders dan een karper of een snoek heeft gevangen? *Do you think it is possible that Patrick also caught something else than a carp or a pike?*

4. (C1/C2) Fleur vertelde aan een vriend dat Sophie vorige week kunst was gaan maken op een kunstcursus, en dat de docent had gezegd dat als **ze een tekening of een collage (C1) / meerdere kunstwerken (C2)** zou maken, ze wat extra hulp zou krijgen. De vriend zei: "Oh, en wat heeft Sophie gemaakt?" Fleur antwoordde: "Ze heeft een tekening of een collage gemaakt."

*4. (C1/C2) Fleur told a friend that Sophie took an art course last week, and that the teacher had said that if she would make **a drawing or a montage (C1) / more than one artwork (C2)**, she would get some extra help. The friend said: "Oh, and what did Sophie make?" Fleur answered: "She made a drawing or a montage."*

4. (C3) Fleur vertelde aan een vriend dat Sophie, Marijne en Linda vorige week kunst waren gaan maken op een kunstcursus, en dat de docent had gezegd dat degene die een tekening of een collage zou maken, wat extra hulp zou krijgen. De vriend zei: "Oh, en wie van hen heeft een tekening of een collage gemaakt?" Fleur antwoordde: "Sophie heeft een tekening of een collage gemaakt."

4. (C3) Fleur told a friend that Sophie, Marijne and Linda took an art course last week, and that the teacher had said that the one who would make a drawing or a montage, would get some extra help. The friend said: "Oh, and who of them made a drawing or a montage?" Fleur answered: "Sophie made a drawing or a montage."

<u>Question-set item 4</u>

CQ (C1):	Wat zou Sophie krijgen als ze een tekening of een collage zou maken?
CQ (C2):	Wat zou Sophie krijgen als ze meerdere kunstwerken zou maken?
CQ (C3+C4):	Wat zouden Sophie, Marijne en Linda krijgen als ze een tekening of een collage zouden maken?
IgnI-Q:	Denk je dat Fleur precies weet wat Sophie heeft gemaakt? *Do you think Fleur knows exactly what Sophie made?*
SI-Q:	Denk je dat het mogelijk is dat Sophie zowel een tekening als een collage heeft gemaakt? *Do you think it is possible that Sophie made both a drawing and a montage?*
Exh-Q:	Denk je dat het mogelijk is dat Sophie ook nog iets anders dan een tekening of een collage heeft gemaakt? *Do you think it is possible that Sophie also made something else than a drawing or a montage?*

5. (C1/C2) Desiree vertelde aan haar moeder dat Lennart gisteren op zijn werk elektronische apparaten was gaan repareren, en dat de baas hem had beloofd dat als hij **een TV of een radio (C1) / minstens twee apparaten (C2)** zou repareren, hij een dag vrij zou krijgen. Desirees moeder zei: "Oh, en wat heeft Lennart gerepareerd?" Desiree antwoordde: "Hij heeft een TV of een radio gerepareerd."

*5. (C1/C2) Desiree told her mother that Lennart was fixing electronic appliances at his work yesterday, and that the boss had promised him that if he would fix **a TV or a radio (C1) / at least two appliances (C2)**, he would get a day off. Desiree's mother said: "Oh, and what did Lennart fix?" Desiree answered: "He fixed a TV or a radio."*

5. (C3) Desiree vertelde aan haar moeder dat Jaap en Lennart gisteren op hun werk elektronische apparaten waren gaan repareren, en dat de baas had beloofd dat degene die een TV of een radio zou repareren, een dag vrij zou krijgen. Desirees moeder zei: "Oh, en wie van hen heeft een TV of een radio gerepareerd? Desiree antwoordde: "Lennart heeft een TV of een radio gerepareerd."

5. (C3) Desiree told her mother that Jaap and Lennart were fixing electronic appliances at their work yesterday, and that the boss had promised them that the one who would fix a TV or a radio, would get a day off. Desiree's mother said: "Oh, and who of them fixed a TV or a radio?" Desiree answered: "Lennart fixed a TV or a radio."

<u>Question-set item 5</u>

CQ (C1+C2): Wat moest Lennart repareren om een dag vrij te krijgen?
CQ (C3+C4): Wat moesten Jaap en Lennart repareren om een dag vrij te krijgen?

IgnI-Q:	Denk je dat Desiree precies weet wat Lennart heeft gerepareerd? *Do you think Desiree knows exactly what Lennart fixed?*
SI-Q:	Denk je dat het mogelijk is dat Lennart zowel een TV als een radio heeft gerepareerd? *Do you think it is possible that Lennart fixed both a TV and a radio?*
Exh-Q:	Denk je dat het mogelijk is dat Lennart ook nog iets anders dan een TV of een radio heeft gerepareerd? *Do you think it is possible that Lennart also fixed something else than a TV or a radio?*

6. (C1/C2) Anouk vertelde aan haar moeder dat Lisa gisteren naar een kledinginzameling voor arme kinderen was geweest, en dat de organisator had gezegd dat als ze **een trui of een jas (C1) / meerdere kledingstukken (C2)** zou schenken, ze een aandenken zou krijgen. Anouks moeder zei: "Oh, en wat heeft Lisa geschonken?" Anouk antwoordde: "Ze heeft een trui of een jas geschonken."

*6. (C1/C2) Anouk told her mother that Lisa went to an event where they collected second hand clothes for poor children, and that the organizer had said that if she would donate **a sweater or a coat (C1) / more than one piece of clothing (C2)**, she would get a souvenir. Anouk's mother said: "Oh, and what did Lisa donate?" Anouk answered: "She donated a sweater or a coat."*

6. (C3) Anouk vertelde aan haar moeder dat Esther en Lisa gisteren naar een kledinginzameling voor arme kinderen waren geweest, en dat de organisator had gezegd dat degene die een trui of een jas zou schenken, een aandenken zouden krijgen. Anouks moeder zei: "Oh, en wie van hen heeft een trui of een jas geschonken?" Anouk antwoordde: "Lisa heeft een trui of een jas geschonken."

6. (C3) Anouk told her mother that Esther and Lisa went to an event where they collected second hand clothes for poor children, and that the organizer had said that if they would donate a sweater or a coat, they would get a souvenir. Anouk's mother said: "Oh, and who of them donated a sweater or a coat?" Anouk answered: "Lisa donated a sweater or a coat."

<u>Question-set item 6</u>

CQ (C1+C2): Wat moest Lisa schenken om een aandenken te krijgen?
CQ (C3+C4): Wat moesten Esther en Lisa schenken om een aandenken te krijgen?

IgnI-Q: Denk je dat Anouk precies weet wat Lisa heeft geschonken?
Do you think Anouk knows exactly what Lisa donated?

SI-Q: Denk je dat het mogelijk is dat Lisa zowel een trui als een jas heeft geschonken?
Do you think it is possible that Lisa donated both a sweater and a coat?

Exh-Q: Denk je dat het mogelijk is dat Lisa ook nog iets anders heeft geschonken dan een trui of een jas?
Do you think it is possible that Lisa also donated something else than a sweater or a coat?

7. (C1/C2) Sanne vertelde aan een vriendin dat Ingrid gisteren naar de Ikea was gegaan, en dat haar huisgenootje had gezegd dat als ze daar **een bank of een stoel (C1) / meer dan één meubelstuk (C2)** zou kopen, ze echt gek was. De vriendin zei: "Oh, en wat heeft Ingrid gekocht bij de Ikea?" Sanne antwoordde: "Ze heeft een bank of een stoel gekocht."

*7. (C1/C2) Sanne told a friend (fem.) that Ingrid went to Ikea yesterday, and that her housemate had said that if she would buy **a couch or a chair (C1) / more than one piece of furniture (C2)** there, she really was out of her mind. The friend said: "Oh, and what did Ingrid buy at Ikea?" Sanne answered: "She bought a couch or a chair."*

7. (C3) Sanne vertelde aan een vriendin dat Ingrid en Marjan gisteren naar de Ikea waren gegaan, en dat hun huisgenootje had gezegd dat degene die daar een bank of een stoel zou kopen, echt gek was. De vriendin zei: "Oh, en wie van hen heeft een bank of een stoel gekocht?" Sanne antwoordde: "Ingrid heeft een bank of een stoel gekocht."

Question-set item 7

CQ (C1):	Wat was Ingrid volgens haar huisgenootje als ze een bank of een stoel zou kopen?
CQ (C2):	Wat was Ingrid volgens haar huisgenootje als ze meer dan één meubelstuk zou kopen?
CQ (C3+C4):	Wat waren Ingrid en Marjan volgens hun huisgenootje als ze een bank of een stoel zouden kopen?
IgnI-Q:	Denk je dat Sanne precies weet wat Ingrid heeft gekocht?
SI-Q:	Denk je dat het mogelijk is dat Ingrid zowel een bank als een stoel heeft gekocht?
Exh-Q:	Denk je dat het mogelijk is dat Ingrid ook nog iets anders dan een bank of een stoel heeft gekocht?

8. (C1/C2) Carolien vertelde aan haar moeder dat Yvette gisteren had meegedaan aan een muziekworkshop met allerlei instrumenten, en dat de docent had gezegd dat als ze **een saxofoon of een trompet (C1) / meerder instrumenten (C2)** zou uitproberen, ze applaus van hem zou krijgen. Caroliens moeder zei: "Oh, en wat heeft Yvette uitgeprobeerd?" Carolien antwoordde: "Ze heeft een saxofoon of een trompet uitgeprobeerd."

*8. (C1/C2) Carolien told her mother that Yvette participated in a music workshop yesterday with all kinds of musical instruments, and that the teacher had said that if she would try out **a saxophone or a trumpet (C1) / more than one musical instrument (C2)**, he would applaud. Carolien's mother said: "Oh, and what did Yvette try out?" Carolien answered: "She tried out a saxophone or a trumpet."*

8. (C3) Carolien vertelde aan haar moeder dat Yvette en Anne gisteren hadden meegedaan aan een muziekworkshop met allerlei instrumenten, en dat de docent had gezegd dat degene die een saxofoon of een trompet zou uitproberen, applaus van hem zou krijgen. Caroliens moeder zei: "Oh, en wie van hen heeft een saxofoon of een trompet uitgeprobeerd?" Carolien antwoordde: "Yvette heeft een saxofoon of een trompet uitgeprobeerd."

Question-set item 8
CQ (C1+C2): Wat moest Yvette uitproberen om applaus te krijgen?
CQ (C3+C4): Wat moesten Yvette en Anne uitproberen om applaus te krijgen?
IgnI-Q: Denk je dat Carolien precies weet wat Yvette heeft uitgeprobeerd?
SI-Q: Denk je dat het mogelijk is dat Yvette zowel een saxofoon als een trompet heeft uitgeprobeerd?
Exh-Q: Denk je dat het mogelijk is dat Yvette ook nog iets anders dan een saxofoon of een trompet heeft uitgeprobeerd?

9. (C1/C2) Martine vertelde aan haar tante dat Wendy gisteren bij handenarbeid op school zelf kleding was gaan maken, en dat de docent had gezegd dat als ze **een broek of een blouse (C1) / meerdere kledingstukken (C2)** zou maken, ze een extra hoog cijfer zou krijgen. De tante zei: "Oh, en wat heeft Wendy gemaakt?" Martine antwoordde: "Ze heeft een broek of een blouse gemaakt."

*9. (C1/C2) Martine told her aunt that Wendy was making clothes for handicraft class at school yesterday, and that the teacher said that if she would make **pants or a blouse (C1) / more than one piece of clothing (C2)**, she would get a higher grade. The aunt said: "Oh, and what did Wendy make?" Martine answered: "She made pants or a blouse."*

9. (C3) Martine vertelde aan haar tante dat Wendy en Ingrid gisteren bij handenarbeid op school allebei zelf kleding waren gaan maken, en dat de docent had gezegd dat degene die een broek of een blouse zou maken, een extra hoog cijfer zou krijgen. De tante zei: "Oh, en wie van hen heeft een broek of een blouse gemaakt?" Martine antwoordde: "Wendy heeft een broek of een blouse gemaakt."

Question-set item 9
CQ (C1+C2): Wat moest Wendy maken om een extra hoog cijfer te krijgen?
CQ (C3+C4): Wat moesten Wendy en Ingrid maken om een extra hoog cijfer te krijgen?
IgnI-Q: Denk je dat Martine precies weet wat Wendy heeft gemaakt?
SI-Q: Denk je dat het mogelijk is dat Wendy zowel een broek als een blouse heeft gemaakt?
Exh-Q: Denk je dat het mogelijk is dat Wendy ook nog iets anders heeft gemaakt dan een broek of een blouse?

10. (C1/C2) Yvonne vertelde aan een vriendin dat Richard gisteren naar een beurs voor computeronderdelen was geweest, en dat als je daar **een toetsenbord of een muis (C1) / meerdere onderdelen (C2)** kocht, je er een gratis artikel bij kreeg. De vriendin zei: "Oh, en wat heeft Richard gekocht?" Yvonne antwoordde: "Hij heeft een toetsenbord of een muis gekocht."

*10. (C1/C2) Yvonne told a friend (fem.) that Richard went to a fair for computer parts, and if one bought **a keyboard or a mouse (C1) / more than one computer part (C2)** there, one got an extra item for free. The friend said: "Oh, and what did Richard buy?" Yvonne answered: He bought a keyboard or a mouse."*

10. (C3) Yvonne vertelde aan een vriendin dat Richard en Wilco gisteren naar een beurs voor computeronderdelen waren geweest, en dat als je daar een toetsenbord of een muis kocht, je er een gratis artikel bij kreeg. De vriendin zei: "Oh, en wie van hen heeft een toetsenbord of een muis gekocht?" Yvonne antwoordde: "Richard heeft een toetsenbord of een muis gekocht op de beurs."

Question-set item 10

CQ (C1): Wat zou Richard krijgen als hij een toetsenbord of een muis zou kopen?
CQ (C2): Wat zou Richard krijgen als hij meerdere onderdelen zou kopen?
CQ (C3+C4): Wat zouden Richard en Wilco krijgen als ze een toetsenbord of een muis zouden kopen?
IgnI-Q: Denk je dat Yvonne precies weet wat Richard heeft gekocht?
SI-Q: Denk je dat het mogelijk is dat Richard zowel een toetsenbord als een muis heeft gekocht?
Exh-Q: Denk je dat het mogelijk is dat Richard ook nog iets anders dan een toetsenbord of een muis heeft gekocht?

11. (C1/C2) Anita vertelde aan haar moeder dat Bart op vakantie enge beesten was gaan vangen in de jungle, en dat de reisleider hem een beloning had beloofd als hij **een spin of een slang (C1) / meer dan één eng beest (C2)** zou vangen. De moeder zei: "Oh, en wat heeft Bart gevangen?" Anita antwoordde: "Hij heeft een spin of een slang gevangen."

*11. (C1/C2) Anita told her mother that Bart went out to catch creepy animals in the jungle during his vacation, and that the tour guide had promised him a reward if he would catch **a spider or a snake (C1) / more than one creepy animal (C2)**. The mother said: "Oh, and what did Bart catch?" Anita answered: "He caught a spider or a snake."*

11. (C3) Anita vertelde aan haar moeder dat Bart en Karel op vakantie enge beesten waren gaan vangen in de jungle, en dat de reisleider een beloning had beloofd aan degene die een spin of een slang zou vangen. De moeder zei: "Oh, en wie van hen heeft een spin of een slang gevangen?" Anita antwoordde: "Bart heeft een spin of een slang gevangen."

Question-set item 11

CQ (C1): Wie had aan Bart een beloning beloofd als hij een spin of een slang zou vangen?
CQ (C2): Wie had aan Bart een beloning beloofd als hij meer dan één eng beest zou vangen?
CQ (C3+C4): Wie had een beloning beloofd aan degene die een spin of een slang zou vangen?
IgnI-Q: Denk je dat Anita precies weet wat Bart heeft gevangen?
SI-Q: Denk je dat het mogelijk is dat Bart zowel een spin als een slang heeft gevangen?
Exh-Q: Denk je dat het mogelijk is dat Bart ook nog iets anders dan een spin of een schorpioen had gevangen?

12. (C1/C2) Cynthia vertelde aan haar zus dat Sharon gisteren planten was gaan kopen bij het tuincentrum, en dat haar huisgenoot had gezegd dat als ze **een varen of een geranium (C1) / meerder planten (C2)** zou kopen, hij haar keihard uit zou lachen. De zus zei: "Oh, en wat heeft Sharon gekocht?" Cynthia antwoordde: "Ze heeft een varen of een geranium gekocht."

*12. (C1/C2) Cynthia told her sister that Sharon went to buy plants yesterday at the garden center, and that her housemate had said that if she would buy **a fern or a geranium (C1) / more than one plant (C2)**, he would laugh at her. The sister said: "Oh, and what did Sharon buy?" Cynthia answered: "She bought a fern or a geranium."*

12. (C3) Cynthia vertelde aan haar zus dat Sharon en Rachel gisteren planten waren gaan kopen bij het tuincentrum, en dat hun huisgenoot had gezegd dat hij degene die een varen of een geranium zou kopen, keihard uit zou lachen. De zus zei: "Oh, en wie van hen heeft een varen of een geranium gekocht?" Cynthia antwoordde: "Sharon heeft een varen of een geranium gekocht."

Question-set item 12

CQ (C1): Wie had gezegd dat hij Sharon keihard uit zou lachen als ze een varen of een geranium zou kopen?
CQ (C2): Wie had gezegd dat hij Sharon keihard uit zou lachen als ze meerdere planten zou kopen?
CQ (C3+C4): Wie had gezegd dat hij degene die een varen of een geranium zou kopen keihard uit zou lachen?
IgnI-Q: Denk je dat Cynthia precies weet wat Sharon heeft gekocht?
SI-Q: Denk je dat het mogelijk is dat Sharon zowel een varen als een geranium heeft gekocht?
Exh-Q: Denk je dat het mogelijk is dat Sharon ook nog iets anders dan een varen of een geranium heeft gekocht?

Exp 4 Average IgnI- and SI-rates per item (percentages)

	C1 [−alt, +foc]		C2 [+alt, +foc]		C3 [−alt, −foc]		C4 [control]	
	IgnI	SI	IgnI	SI	IgnI	SI	IgnI	SI
1	90	40	100	85	91	64	50	25
2	94	75	90	80	92	69	36	27
3	94	75	100	100	100	77	73	18
4	90	50	100	92	91	73	31	31
5	91	45	100	44	100	30	62	31
6	100	64	94	75	100	60	54	46
7	92	85	100	64	88	50	40	20
8	92	77	91	27	81	69	70	30
9	92	77	91	55	81	88	30	30
10	100	50	92	69	100	36	50	19
11	88	69	100	90	92	85	55	18
12	91	73	94	88	100	30	23	23

Exp 6 Average SI- and exh-rates per item (percentages)

	C1 [-alt, +foc]		C2 [+alt, +foc]		C3 [-alt, -foc]		C4 [control]	
	SI	exh	SI	exh	SI	exh	SI	exh
1	56	44	87	50	56	17	25	0
2	75	60	75	89	80	10	38	0
3	81	10	94	89	73	10	25	0
4	69	56	87	70	81	67	25	30
5	56	50	56	30	63	0	20	10
6	44	50	69	70	63	33	47	30
7	80	40	60	50	63	0	31	0
8	60	50	69	83	56	30	19	0
9	87	67	63	50	88	56	69	33
10	63	22	87	90	44	0	13	20
11	81	10	69	67	73	10	13	33
12	63	83	81	60	63	0	27	0

APPENDIX 5

Items of Experiments 5 and 7 (chapters 4 and 9)

(C1) = [−alternative relevance, +focus]
(C2) = [+alternative relevance, +focus]
(C3) = [−alternative relevance, −focus]
(C4) = control

In both experiments, all items contained two questions, the first was the comprehension question (CQ) (which varied over conditions) and the second was either the scalar implicature question (SI-Q) or the external exhaustivity question (Exh-Q). (These labels were absent in the actual items).

Items 1-8 always contained an SI-question, and items 9-16 always contained an external exhaustivity question. Items 1-12 were based on the items of Experiments 4 and 6 (see Appendix 4) and items 13-16 were new. Therefore, I only provide translations for items 13-16 (and item 1 for clarity). Translations of items 1-12 can also be found in Appendix 7.

I present the items of all conditions together, and highlight the differences in boldface (not in the actual items). Condition 4 is again identical to condition 3 but with a term answer instead of a sentential answer.

Items start on the next page.

1. Marieke vertelde aan een vriendin dat Laura en Barbara gisteren zeedieren waren gaan zoeken op het strand, en dat toen ze terugkwamen, hun moeder zei dat als één van hen **een krab of een mossel (C1+C3) / minstens twee zeedieren (C2)** had gevonden op het strand, diegene die avond later mocht opblijven.
De vriendin zei: **"Oh, en wat had Laura gevonden op het strand?" (C1+C2) / "Oh, en wie van Laura en Barbara had een krab of een mossel gevonden op het strand?" (C3)**
Marieke antwoordde: "Laura had een krab of een mossel gevonden op het strand."

CQ (C1+C3): Wat zou degene mogen die een krab of een mossel had gevonden?
CQ (C2): Wat zou degene mogen die minstens twee zeedieren had gevonden?
SI-Q: Denk je dat het mogelijk is dat Laura zowel een krab als een mossel had gevonden?

1. Marieke told a friend that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found ***a crab or a mussel (C1+C3) / at least two marine animals (C2)*** *on the beach, that person got to stay up late that night.*
The friend said: ***"Oh, and what had Laura found on the beach?" (C1+C2) / "Oh, and who of Laura and Barbara had found a crab or a mussel on the beach?" (C3)***
Marieke answered: "Laura had found a crab or a mussel on the beach."

CQ (C1+C3): What would the person who found a crab or a mussel be allowed to do?
CQ (C2): What would the person who found at least two marine animals be allowed to do?
SI-Q: Do you think it is possible that Laura had found both a crab and a mussel?

2. Frans vertelde aan een vriend dat Chris en Marco gisteren naar de snackbar waren geweest, maar dat toen ze terugkwamen, hun vader zei dat als één van hen **een kroket of een frikadel (C1+C3) / meer dan één snack (C2)** had genomen bij de snackbar, diegene geen avondeten kreeg.
De vriend zei: **"Oh, en wat had Chris genomen bij de snackbar?" (C1+C2) / "Oh, en wie van Chris en Marco had een kroket of een frikadel genomen bij de snackbar?" (C3)**
Frans antwoordde: "Chris had een kroket of een frikadel genomen bij de snackbar."
CQ (C1+C3): Wie zei dat degene die een kroket of een frikadel had genomen, geen avondeten kreeg?
CQ (C2): Wie zei dat degene die meer dan één snack had genomen, geen avondeten kreeg?
SI-Q: Denk je dat het mogelijk is dat Chris zowel een kroket als een frikadel had genomen?

3. Guido vertelde aan een vriend dat Patrick en Maurice gisteren waren gaan vissen bij het meer, en dat hun vader na afloop zei dat als één van hen **een karper of een snoek (C1+C3) / hen minstens twee vissen (C2)** had gevangen bij het meer, diegene een tientje kreeg.
De vriend zei: **"Oh, en wat had Patrick gevangen bij het meer?" (C1+C2) / "Oh, en wie van Patrick en Maurice had een karper of een snoek gevangen bij het meer?" (C3)**
Guido antwoordde: "Patrick had een karper of een snoek gevangen bij het meer."

CQ (C1+C3): Wat moesten Patrick en Maurice gevangen hebben om een tientje te krijgen?
CQ (C2): Wat moesten Patrick en Maurice gevangen hebben om een tientje te krijgen?
SI-Q: Denk je dat het mogelijk is dat Patrick zowel een karper als een snoek had gevangen?

4. Fleur vertelde aan een vriendin dat Marijne en Sophie vorige week kunst waren gaan maken op een kunstcursus, en dat de docente aan het einde zei dat degene die **een ets of een collage (C1+C3) / meerdere kunstwerken (C2)** had gemaakt op de cursus, extra materiaalkosten moest betalen.
De vriendin zei: **"Oh, en wat had Sophie gemaakt op de cursus?" (C1+C2) / "Oh, en wie van Marijne en Sophie had een ets of een collage gemaakt op de cursus?" (C3)**
Fleur antwoordde: "Sophie had een ets of een collage gemaakt op de cursus."

CQ (C1+C3): Wat zou degene moeten die een ets of een collage had gemaakt?
CQ (C2): Wat zou degene moeten die meerdere kunstwerken had gemaakt?
SI-Q: Denk je dat het mogelijk is dat Sophie zowel een ets als een collage had gemaakt?

5. Nick vertelde aan een vriend dat Jaap en Lennart gisteren in de werkplaats elektronische apparaten waren gaan repareren, en dat aan het einde van de dag de baas zei dat degene die **een TV of een radio (C1+C3) / minstens twee apparaten (C2)** had gerepareerd in de werkplaats, de dag erna vrij kreeg.
De vriend zei: **"Oh, en wat had Lennart gerepareerd in de werkplaats?" (C1+C2) / "Oh, en wie van Jaap en Lennart had een TV of een radio gerepareerd in de werkplaats?" (C3)**
Nick antwoordde: "Lennart had een TV of een radio gerepareerd in de werkplaats."

CQ (C1+C3): Wat zou degene krijgen die een TV of een radio had gerepareerd?
CQ (C2) Wat zou degene krijgen die minstens twee apparaten had gerepareerd?
SI-Q: Denk je dat het mogelijk is dat Lennart zowel een TV als een radio had gerepareerd?

6. Anouk vertelde aan een vriendin dat Saskia en Lisa gisteren naar een kledinginzameling voor arme kinderen waren geweest, en dat aan het einde de organisatrice zei dat als één van hen **een trui of een jas (C1+C3) / meerdere kledingstukken (C2)** had geschonken bij de inzameling, diegene een aandenken kreeg.
De vriendin zei: **"Oh, en wat had Lisa geschonken bij de inzameling?" (C1+C2) / "Oh, en wie van Saskia en Lisa had een trui of een jas geschonken bij de inzameling?" (C3)**
Anouk antwoordde: "Lisa had een trui of een jas geschonken bij de inzameling."

CQ (C1+C3): Wat zou degene krijgen die een trui of een jas had geschonken?
CQ (C2): Wat zou degene krijgen die meerdere kledingstukken had geschonken?
SI-Q: Denk je dat het mogelijk is dat Lisa zowel een trui als een jas had geschonken?

7. Sanne vertelde aan een vriendin dat Jenny en Marjan gisteren naar de Ikea waren gegaan, en dat er een actie was dat als je **een bank of een stoel (C1+C3) / meerdere meubels (C2)** had gekocht bij de Ikea, je kans maakte op een prijs.
De vriendin zei: **"Oh, en wat had Jenny gekocht bij de Ikea?" (C1+C2) / "Oh, en wie van Jenny en Marjan had een bank of een stoel gekocht bij de Ikea?" (C3)**
Sanne antwoordde: "Jenny had een bank of een stoel gekocht bij de Ikea."

CQ (C1+C3): Wat moest je gekocht hebben om kans te maken op een prijs?
CQ (C2): Wat moest je gekocht hebben om kans te maken op een prijs?
SI-Q: Denk je dat het mogelijk is dat Jenny zowel een bank als een stoel had gekocht?

8. Tim vertelde aan een vriend dat Sebastiaan en Mark gisteren hadden meegedaan aan een muziekworkshop, en dat de leraar aan het einde zei dat degene die **een saxofoon of een trompet (C1+C3) / meerdere instrumenten (C2)** had uitgeprobeerd bij de workshop, een applaus kreeg.
De vriend zei: **"Oh, en wat had Mark uitgeprobeerd bij de workshop?" (C1+C2) / "Oh, en wie van Sebastiaan en Mark had een saxofoon of een trompet uitgeprobeerd bij de workshop?" (C3)**
Tim antwoordde: "Mark had een saxofoon of een trompet uitgeprobeerd bij de workshop."

CQ (C1+C3): Wie zei dat degene die een saxofoon of een trompet had uitgeprobeerd bij de workshop, een applaus kreeg?
CQ (C2): Wie zei dat degene die meerdere instrumenten had uitgeprobeerd bij de workshop, een applaus kreeg?
SI-Q: Denk je dat het mogelijk is dat Mark zowel een saxofoon als een trompet had uitgeprobeerd?

9. Martine vertelde aan een vriendin dat Wendy en Ingrid gisteren op een cursus kleding waren gaan maken, en dat de lerares aan het einde zei dat degene die **een broek of een blouse (C1+C3) / meerdere kledingstukken (C2)** had gemaakt, een extra hoge beoordeling kreeg.
De vriendin zei: **"Oh, en wat had Wendy gemaakt op de cursus?" (C1+C2) / "Oh, en wie van Wendy en Ingrid had een broek of een blouse gemaakt op de cursus?" (C3)**
Martine antwoordde: "Wendy had een broek of een blouse gemaakt op de cursus."

CQ (C1+C3): Wat moesten Wendy en Ingrid gemaakt hebben om een extra hoge beoordeling te krijgen?
CQ (C2): Wat moesten Wendy en Ingrid gemaakt hebben om een extra hoge beoordeling te krijgen?
Exh-Q: Denk je dat het mogelijk is dat Wendy ook nog iets anders dan een broek of een blouse had gemaakt?

10. Yvonne vertelde aan een vriendin dat Rianne en Nicole gisteren naar een beurs voor computeronderdelen waren geweest, en dat als je **een toetsenbord of een muis (C1+C3) / meerdere artikelen (C2)** had gekocht op de beurs, je een kortingsbon voor een computerwinkel kreeg.
De vriend zei: **"Oh, en wat had Richard gekocht op de beurs?" (C1+C2) / "Oh, en wie van Richard en Wilco had een toetsenbord of een muis gekocht op de beurs?" (C3)**
Peter antwoordde: "Richard had een toetsenbord of een muis gekocht op de beurs."

CQ (C1+C3): Wat moest je gekocht hebben om een kortingsbon voor een computerwinkel te krijgen?
CQ (C2): Wat moest je gekocht hebben om een kortingsbon voor een computerwinkel te krijgen?
Exh-Q: Denk je dat het mogelijk is dat Richard ook nog iets anders dan een toetsenbord of een muis had gekocht op de beurs?

11. Emiel vertelde aan een vriend dat Bart en Karel op vakantie enge beesten waren gaan vangen in de jungle, en dat toen ze terugkwamen, de reisleider zei dat degene die **een spin of een schorpioen (C1+C3) / meer dan één eng beest (C2)** had gevangen in de jungle, een beloning kreeg.
De vriend zei: **"Oh, en wat had Bart gevangen in de jungle?" (C1+C2) / "Oh, en wie van Bart en Karel had een spin of een schorpioen gevangen in de jungle?" (C3)**
Emiel antwoordde: "Bart had een spin of een schorpioen gevangen in de jungle."

CQ (C1+C3): Wie zei dat degene die een spin of een schorpioen had gevangen in de jungle, een beloning kreeg?
CQ (C2) : Wie zei dat degene die meer dan één eng beest had gevangen in de jungle, een beloning kreeg?
Exh-Q: Denk je dat het mogelijk is dat Bart ook nog iets anders dan een spin of een schorpioen had gevangen?

12. Claudia vertelde aan een vriendin dat Paula en Martine gisteren planten waren gaan kopen bij het tuincentrum, en dat er een actie was dat als je **een varen of een geranium (C1+C3) / meerdere planten (C2)** had gekocht bij het tuincentrum, je korting kreeg bij de bouwmarkt ernaast.
De vriendin zei: **"Oh, en wat had Paula gekocht bij het tuincentrum?" (C1+C2) / "Oh, en wie van Paula en Martine had een varen of een geranium gekocht bij het tuincentrum?" (C3)**
Claudia antwoordde: "Paula had een varen of een geranium gekocht bij het tuincentrum."

CQ (C1+C3): Wat moest je gekocht hebben om korting te krijgen bij de bouwmarkt?
CQ (C2): CQ: Wat moest je gekocht hebben om korting te krijgen bij de bouwmarkt?
Exh-Q: Denk je dat het mogelijk is dat Paula ook nog iets anders dan een varen of een geranium had gekocht?

13. Stefan vertelde aan een vriend dat Rob en Floris gisteren naar de kermis waren geweest, en dat bij de uitgang bleek dat degene die **een knuffel of een skippybal (C1+C3) / meer dan één prijs (C2)** had gewonnen op de kermis, gratis op de foto mocht.
De vriend zei: **"Oh, en wat had Rob gewonnen op de kermis?" (C1+C2) / "Oh, en wie van Rob en Floris had een knuffel of een skippybal gewonnen op de kermis?" (C3)**
Stefan antwoordde: "Rob had een knuffel of een skippybal gewonnen op de kermis."

CQ (C1+C3): Wat zou degene mogen die een knuffel of een skippybal had gewonnen?
CQ (C2): Wat zou degene mogen die meer dan één prijs had gewonnen?
Exh-Q: Denk je dat het mogelijk is dat Rob nog iets anders dan een knuffel of een skippybal had gewonnen?

*13. Stefan told a friend that Rob and Floris went to the fair yesterday, and that at the exit it turned out that the one who had won **a teddy bear or a skippy ball (C1+C3) / more than one prize (C2)** at the fair, got to have his picture taken for free.*
*The friend said**: "Oh, and what had Rob won at the fair?" (C1+C2) / "Oh, and who of Rob and Floris had won a teddy bear or a skippy ball at the fair?" (C3)***
The friend said: "Rob had won a teddy bear or a skippy ball at the fair."

CQ (C1+C3): What would the one who won a teddy bear or a skippy ball get?
CQ (C2): What would the one who won more than one prize get?
Exh-Q: Do you think it is possible that Rob also caught something else than a teddy bear or a skippy ball?

14. Ernst vertelde aan een vriend dat Diederick en Roderick gisteren waren gaan jagen, en dat toen ze terugkwamen, de voorzittter van de jachtclub zei dat degene die **een fazant of een konijn (C1+C3) / minstens twee beesten (C2)** had geschoten, lid mocht worden van de jachtclub.
De vriend zei: **"Oh, en wat had Diederick geschoten bij de jacht?" (C1+C2) / "Oh, en wie van Diederick en Roderick had een fazant of een konijn geschoten bij de jacht?" (C3)**
Ernst antwoordde: "Diederick had een fazant of een konijn geschoten bij de jacht."

CQ (C1+C3): Wie zei dat degene die een fazant of een konijn had geschoten, lid mocht worden van de jachtclub?
CQ (C2): Wie zei dat degene die minstens twee beesten had geschoten, lid mocht worden van de jachtclub?
Exh-Q: Denk je dat het mogelijk is dat Diederick ook nog iets anders dan een fazant of een konijn had geschoten?

*14. Ernst told a friend that Diederick and Roderick went hunting for game yesterday, and that when they returned, the president of the hunting club told them that the one who shot **a pheasant or a rabbit (C1+C3) / at least two animals (C2)**, could become a member of the club.*
*The friend said: **"Oh, and what did Diederick shoot at the hunt? (C1+C2) / "Oh, and who of Diederick and Roderick had shot a pheasant or a rabbit at the hunt? (C3)***
Ernst answered: "Diederick shot a pheasant or a rabbit at the hunt."

CQ (C1+C3): Who said that the one who shot a pheasant or a rabbit, could become a member of the club?
CQ (C2): Who said that the one who shot at least two animals, could become a member of the club?
Exh-Q: Do you think it is possible that Diederick had also shot something else than a pheasant or a rabbit?

15. Leonie vertelde aan een vriendin dat Charlotte en Roos gisteren op een archeologische afgraving naar oude resten waren gaan zoeken, en dat de leidster aan het einde van de dag zei dat degene die **een munt of een potscherf (C1+C3) / meerdere resten (C2)** had gevonden bij de afgraving, de volgende dag terug mocht komen.
De vriend zei: **"Oh, en wat had Roos gevonden bij de afgraving?" (C1+C2) / "Oh, en wie van Charlotte en Roos had een munt of een potscherf gevonden bij de afgraving?" (C3)**
Leonie antwoordde: "Roos had een munt of een potscherf gevonden bij de afgraving."

CQ (C1+C3): Wie zei dat degene die een munt of een potscherf had gevonden, de volgende dag terug mocht komen?
CQ: (C2): Wie zei dat degene die meerdere resten had gevonden, de volgende dag terug mocht komen?
Exh-Q: Denk je dat het mogelijk is dat Roos ook nog iets anders dan een munt of een potscherf had gevonden?

15. Leonie told a friend (fem.) that Charlotte and Roos went searching for remnants at an archaeological dig yesterday, and that the leader of the dig at the end of the day said that the one who had found ***a coin or a potsherd (C1+C3) / more than one remnant (C2)*** *at the dig, was allowed to come back the next day.*
The friend said: ***"Oh, and what had Roos found at the dig?" (C1+C2) / "Oh, and who of Charlotte and Roos had found a coin or a potsherd at the dig?" (C3)***
Leonie answered: "Roos had found a coin or a potsherd at the dig."

CQ (C1+C3): Who said that the one who had found a coin or a potsherd was allowed to come back the next day?
CQ (C2): Who said that the one who had found more than one remnant was allowed to come back the next day?
Exh-Q: Do you think it is possible that Roos also found something else than a coin or a potsherd?

16. Sylvia vertelde aan een vriendin dat Dagmar en Femke gisteren naar een beurs waren geweest, en dat toen ze terugkwamen, hun huisgenootje zei dat als één van hen **een pen of een sleutelhanger (C1+C3) / meer dan één relatiegeschenk (C2)** had meegenomen van de beurs, diegene echt sneu was.
De vriendin zei: **"Oh, en wat had Dagmar meegenomen van de beurs?" (C1+C2) / "Oh, en wie van Dagmar en Femke had een pen of een sleutelhanger meegenomen van de beurs?" (C3)**
Sylvia antwoordde: "Dagmar had een pen of een sleutelhanger meegenomen van de beurs."

CQ (C1+C3): Wat was degene die een pen of een sleutelhanger had meegenomen volgens het huisgenootje?
CQ (C2): Wat was degene die meer dan één relatiegeschenk had meegenomen volgens het huisgenootje?
Exh-Q: Denk je dat het mogelijk is dat Dagmar ook nog iets anders dan een pen of een sleutelhanger had meegenomen van de beurs?

16. Sylvia told a friend (fem.) that Dagmar and Femke went to a convention yesterday, and that when they returned, their housemate said that if one of them had brought back ***a pen or a keychain (C1+C3) / more than one promotional gift (C2)*** *from the convention, that person was just sad.*
*The friend said****: "Oh, and what had Dagmar brought back from the convention?" (C1+C2) / "Oh, and who of Dagmar and Femke had brought back a pen or a keychain from the convention?" (C3)***
Sylvia answered: Dagmar had brought back a pen or a keychain from the convention."

CQ (C1+C3): What was the person who brought back a pen or a keychain according to the housemate?
CQ (C2): What was the person who brought back more than one promotional gift according to the housemate?
*Exh-Q: Do you think it is possible that Dagmar had also brought back something else from the convention than a pen or a keychain?***Exp 5 Average SI- and exh-rates per item (percentages)**

	C1 [-alt, +foc]		C2 [+alt, +foc]		C3 [-alt, -foc]		C4 [control]	
	SI	exh	SI	exh	SI	exh	SI	exh
1	76		65		71		12	
2	71		82		71		18	
3	71		82		82		0	
4	76		82		53		47	
5	71		82		76		18	
6	71		76		71		18	
7	65		71		59		18	
8	71		71		35		18	
9		41		76		29		6
10		53		65		0		0
11		41		53		18		6
12		47		47		12		12
13		41		65		12		6
14		65		59		6		12
15		35		65		0		0
16		18		29		6		6

Exp 7 Average SI- and exh-rates per item (percentages)

	C1 [-alt, +foc]		C2 [+alt, +foc]		C3 [-alt, -foc]		C4 [control]	
	SI	exh	SI	exh	SI	exh	SI	exh
1	38		90		78		14	
2	71		75		70		11	
3	57		88		70		11	
4	75		90		100		14	
5	67		100		50		10	
6	67		86		50		0	
7	70		56		71		38	
8	70		44		43		38	
9		30		89		14		50
10		63		90		0		0
11		29		63		60		22
12		67		57		38		10
13		43		63		20		33
14		56		71		38		0
15		60		78		14		25
16		38		90		11		14

APPENDIX 6

Items of Experiment 8 (chapter 6)

C1 = focus *or*, C2 = non-focus *or*
Only the *or* conditions are reported as:
C3 (focus *and*) = C1 with all occurrences of *or* replaced by *and*
C4 (non-focus *and*) = C2 with all occurrences of *or* replaced by *and*

I only translate the target sentences for items 13-24.

1. (C1)
Jacques Goderie en René Mioch/waren/naar de Oscar-uitreiking/geweest./Toen Jacques terugkwam,//
vroeg een vriend/welke filmsterren/René/had/geïnterviewd.//
Jacques antwoordde/dat René/Brad Pitt of George Clooney/had/geïnterviewd./De vriend/zei//
dat hij/dat/behoorlijk gaaf/vond.

1. (C2)
Jacques Goderie en René Mioch/waren/naar de Oscar-uitreiking/geweest./Toen/ze/terugkwamen,//
vroeg een vriend/wie van hen/Brad Pitt of George Clooney/had geïnterviewd.//
Jacques antwoordde/dat René/Brad Pitt of George Clooney/had/geïnterviewd./De vriend/zei//
dat hij/dat wel verwacht had.

1. Jacques Goderie and René Mioch had been to the Oscars. When Jacques (C1) / they (C2) returned, a friend asked which movie stars René had interviewed (C1) / who of them had interviewed Brad Pitt or George Clooney (C2).
Jacques answered / that René / Brad Pitt or George Clooney / had / interviewed./ The friend said that he tought that was pretty cool.

2. (C1)
Katja en Bridget/deden mee/aan Fear Factor./Toen Bridget terugkwam/van de opnames,//
vroeg/de BNN-directeur/wat voor ranzigs/Katja/had opgegeten.//
Bridget antwoordde/dat Katja/een tor of een sprinkhaan/had/opgegeten./De BNN-directeur/zei//
dat hij/dat echt/ranzig vond.

2. (C2)
Katja en Bridget/deden mee/aan Fear Factor./Toen ze/terugkwamen/van de opnames,//
vroeg de BNN-directeur/wie van hen/een tor of een sprinkhaan/had opgegeten.//
Bridget antwoordde/dat Katja/een tor of een sprinkhaan/had/opgegeten./De BNN-directeur/zei//
dat hij/Katja een heldin vond.

2. Katja and Bridget were candidates on Fear Factor. When Bridget (C1) / they (C2) got back from the show, their CEO asked what gross stuff Katja had eaten (C1) / who of them had eaten a beetle or a grasshopper (C2).
Bridget answered / that Katja / a beetle or a grasshopper / had / eaten./ The CEO said he thought that was really gross.

3. (C1)
Bassie en Adriaan/gaven/een voorstelling/in het circus./Toen Bassie terugkwam/in de caravan,//
vroeg Robin/welke circusdieren/Adriaan/had bereden.//
Bassie antwoordde/dat Adriaan/een paard of een olifant/had/bereden./Robin/zei//
dat hij/circusdieren/te gek vond.

3. (C2)
Bassie en Adriaan/gaven/een voorstelling/in het circus./Toen ze terugkwamen/in de caravan,//
vroeg Robin/wie van hen/een paard of een olifant/had bereden.//
Bassie antwoordde/dat Adriaan/een paard of een olifant/had/bereden./Robin/zei//
dat hij Adriaan/een geweldige acrobaat/vond.

3. Bassie and Adriaan gave a show in the circus. When Bassie (C1) / they (C2) returned to the trailer, Robin asked which circus animals Adriaan had ridden (C1) / who of them had ridden a horse or an elephant (C2).
Bassie answered / that Adriaan / a horse or an elephant / had / ridden. / Robin said that he loved circus animals.

4. (C1)
Gerard en Gordon/waren/aan het jagen/in het bos./Toen Gerard/terugkwam,//
vroeg René/welke dieren/Gordon/had geschoten.//
Gerard antwoordde/dat Gordon/een eend of een konijn/had/geschoten./René/zei//
dat hij/jagen/eigenlijk slecht/vond.

4. (C2)
Gerard en Gordon/waren/aan het jagen in het bos./Toen ze/terugkwamen,//
vroeg René/wie van hen/een eend of een konijn/had/geschoten.//
Gerard antwoordde/dat Gordon/een eend of een konijn/had/geschoten./René/zei//
dat hij/niet anders had verwacht/van Gordon.

4. Gerard and Gordon were hunting in the forest. When Gerard (C1) / they (C2) returned, René asked which animals Gordon had shot (C1) / who of them had shot a duck or a rabbit (C2).
Gerard answered / that Gordon / a duck or a rabbit / had / shot /. René said he thought hunting was bad.

5. (C1)
Mitchel en Barrie/gingen/naar de supermarkt./Toen Barrie terugkwam,//
vroeg/een klasgenoot/welke dingen/Mitchel/had/gestolen.//
Barrie antwoordde/dat Mitchel/een zak chips of een zak snoep/had/gestolen./De klasgenoot/zei//
dat hij/een goede keuze/had gemaakt.

5. (C2)
Mitchel en Barrie/gingen/naar de supermarkt./Toen ze/terugkwamen,//
vroeg/een klasgenoot/wie van hen/een zak chips of een zak snoep/had gestolen.//
Barrie antwoordde/dat Mitchel/een zak chips of een zak snoep/had/gestolen./De klasgenoot/zei//
dat hij/dat al vermoedde.

5. Mitchel and Barrie went to the grocery store. When Barrie (C1) / they (C2) returned, a classmate asked which things Mitchel had stolen (C1) / who of them stole a bag of potato chips or a bag of candy (C2).
Barrie answered / that Mitchel / a bag of potato chips or a bag of candy / had / stolen /. The classmate said that he had made a good choice.

6. (C1)
Fokke en Sukke/gingen/naar het eerste college/van een cursus./Toen Fokke aankwam,//
vroeg de docent/welke lesmaterialen/Sukke/had/gekocht.//
Fokke antwoordde/dat Sukke/het boek of de reader/had/gekocht./De docent/zei//
dat de lesmaterialen/achterhaald/waren.

6. (C2)
Fokke en Sukke/gingen/naar het eerste college/van een cursus./Toen ze aankwamen,//
vroeg de docent/wie van hen/het boek of de reader/had/gekocht.//
Fokke antwoordde/dat Sukke/het boek of de reader/had/gekocht./De docent/zei//
dat iedereen/het boek moest hebben.

6. Fokke and Sukke went to the first class of a college course. When Fokke (C1) / they (C2) arrived, the professor asked which course materials Sukke had bought (C1) / who of them bought the book or the reader (C2).
Fokke answered / that Sukke / the book or the reader / had bought. / The teacher said that the course materials were outdated.

7. (C1)
Claus en Bernhard/gingen/naar de Media Markt./Toen Bernhard/terugkwam,//
vroeg/Juliana/welke dingen/Claus/had gekocht.// Bernhard antwoordde/dat Claus/een cd of een dvd/had/gekocht./Juliana/zei// dat ze/nog nooit/van die dingen had gehoord.

7. (C2)
Claus en Bernhard/gingen/naar de Media Markt./Toen ze terugkwamen,//
vroeg/Juliana/wie van hen/een cd of een dvd/had/gekocht.//
Bernhard antwoordde/dat Claus/een cd of een dvd/had/gekocht./Juliana/zei//
dat ze/dat wel verwacht had/van Claus.

7. Claus and Berhard went to an electronics store. When Bernhard (C1) / they (C2) returned, Juliana asked which things Claus had bought (C1) / who of them had bought a cd or a dvd (C2).
Berhard answered / that Claus / a cd or a dvd / had / bought /. Juliana said she had expected that of Claus.

8. (C1)
Dunya en Desie/gingen/winkelen/in de stad./Toen Dunya terugkwam,//
vroeg/een vriendin/welke dingen/Desie/had gekocht.//
Dunya antwoordde/dat Desie/kleding of schoenen/had/gekocht./De vriendin/zei//
dat ze/daar/erg verbaasd/over/was.

8. (C2)
Dunya en Desie/gingen/winkelen/in de stad./Toen ze terugkwamen,//
vroeg/een vriendin/wie van hen/kleding of schoenen/had gekocht.//
Dunya antwoordde/dat Desie/kleding of schoenen/had/gekocht./De vriendin/zei//
dat ze/dat al dacht.

8. Dunya and Desie went shopping in the city. When Dunya (C1) / they (C2) returned, a friend asked which things Desie had bought (C1) / who of them had bought clothing or shoes (C2).
Dunya answered / that Desie / clothing or shoes / had / bought./ The friend said she was very surprised about that.

9. (C1)
Eva en Maria/waren aan het werk/in een naaiatelier./Toen Eva wegging,//
vroeg/hun baas/wat voor kleding/Maria/die dag had/afgemaakt.//
Eva antwoordde/dat Maria/een rok of een broek/had/afgemaakt./De baas/zei//
dat/dat/niet de bedoeling/was.

9. (C2)
Eva en Maria/waren/aan het werk/in een naaiatelier./Toen ze weggingen,//
vroeg hun baas/wie van hen die dag/een rok of een broek/had/afgemaakt.//
Eva antwoordde/dat Maria/een rok of een broek/had/afgemaakt./De baas/zei//
dat hij/trots op Maria/was.

9. Eva and Marie were working in a sweatshop. When Eva (C1) / they (C2) left, they boss asked what clothes Maria had finished that day (C1) / who of them had finished a skirt or pants that day (C2).
Eva answered / that Maria / a skirt or pants / had / finished. /The boss said that he was proud of her.

10. (C1)
Hielke en Sietse/waren aan het vissen/op/hun boot./Toen Hielke thuiskwam,//
vroeg/hun vader/wat voor vissen/Sietse/had gevangen.//
Hielke antwoordde/dat Sietse/een snoek of een karper/had/gevangen./De vader/zei//
dat/dat/een mooie vangst was.

10. (C2)
Hielke en Sietse/waren/aan het vissen/op hun boot./Toen ze/thuiskwamen,//
vroeg hun vader/wie van hen/een snoek of een karper/had gevangen.//
Hielke antwoordde/dat Sietse/een snoek of een karper/had/gevangen./Vader/zei//
dat Sietse/een geluksvogel was.

10. Hielke and Sietse were fishing on their boat. When Hielke (C1) / they (C2) returned, their father asked what kind of fish Sietse had caught (C1) / who of them had caught a pike or a carp (C2).
Hielke answered / that Sietse / a pike or a carp / had caught. / Father said that that was a nice catch.

11. (C1)
Ronald en Frank de Boer/waren jarig./Toen Ronald/op de training/kwam,//
vroeg een teamgenoot/wat voor cadeaus/Frank/had/gekregen.//
Ronald antwoordde/dat Frank/een boekenbon of een bioscoopbon/had/gekregen./De teamgenoot/zei//
dat/dat/weinig origineel/was.

11. (C2)
Ronald en Frank de Boer/waren jarig./Toen ze/op de training kwamen,//
vroeg/een teamgenoot/wie van hen/een boekenbon of een bioscoopbon/had/gekregen.//
Ronald antwoordde/dat Frank/een boekenbon of een bioscoopbon/had/gekregen./De teamgenoot/zei// dat/dat wel bij Frank/paste.

11. Ronald and Frank de Boer were having their birthday. When Ronald (C1) / they (C2) arrived at soccer practice, a teammate asked what kind of gifts Frank had gotten (C1) / who of them had gotten a book gift certificate or a movie gift certificate (C2).
Ronald answered/ that Frank / a book gift certificate or a movie gift certificate / had / gotten. The teammate said that that was not very original.

12. (C1)
Edgar en Jos waren/aan het werk/in een meubelmakerij./Toen Jos/wegging,//
vroeg/de baas/wat voor meubels/Edgar/gemaakt had.//
Jos antwoordde/dat Edgar/een kast of een stoel/had/gemaakt./De baas/zei//
dat hij/daar/geen opdracht/voor had gegeven.

12. (C2)
Edgar en Jos/waren/aan het werk/in een meubelmakerij./Toen ze weggingen,//
vroeg de baas/wie van hen/een kast of een stoel/had/gemaakt.//
Jos antwoordde/dat Edgar/een kast of een stoel/had/gemaakt./De baas/zei//
dat Edgar/een goede meubelmaker was.

12. Edgar and Jos were werking in the furniture-making shop. When Jos (C1) / they (C2) left, the boss asked which pieces of furniture Edgar had made (C1) / who of them had made a cupboard or a chair (C2).
Jos answered / that Edgar /a cupboard or a chair / had / made. / The boss said Edgar was a good craftsman.

13. (C1)
Sylvie en Rafael/deden/mee/aan een muziekworkshop./Toen Sylvie terugkwam,//
vroeg/hun huishoudster/wat voor instrumenten/Rafael/had bespeeld.//
Sylvie antwoordde/dat Rafael/een trommel of een fluit/had/bespeeld./De huishoudster/zei//
dat ze/al moest lachen/bij het idee.

13. (C2)
Sylvie en Rafael/deden mee aan/een muziekworkshop./Toen ze/terugkwamen,//
vroeg hun huishoudster/wie van hen/een trommel of een fluit/had/bespeeld.//
Sylvie antwoordde/dat Rafael/een trommel of een fluit/had/bespeeld./De huishoudster/zei//
dat ze/dat echt/iets voor Rafael/vond.

13. Sylvie answered / that Rafael / a drum or a flute / had / played.

14. (C1)
Diederik en Roderick/gingen/naar de sociëteit./Toen Diederik terugkwam,//
vroegen/zijn huisgenoten/wat voor drank/Roderick/had/gedronken.//
Diederik antwoordde/dat Roderick/port of sterke drank/had/gedronken./De huisgenoten/zeiden//
dat dat/misschien/niet zo slim/was.

14. (C2)
Diederik en Roderick/gingen/naar de sociëteit./Toen ze/terugkwamen,//
vroegen/hun huisgenoten/wie van hen/port of sterke drank/had gedronken.//
Diederik antwoordde/dat Roderick/port of sterke drank/had/gedronken./De huisgenoten/zeiden//
dat Roderick/een waar drankorgel/was.

14. Diederick answered / that Roderick / port or booze / had / drunk.

15. (C1)
Bush en Clinton/hielden/een toespraak/in het Congres./Toen Bush naar buiten kwam,//
vroeg een journalist/welke conflicten/Clinton/had/genoemd.//
Bush antwoordde/dat Clinton/Irak of Afghanistan/had/genoemd./De journalist/zei//
dat/hij/dat dapper vond.

15. (C2)
Bush en Clinton/hielden een toespraak/in het Congres./Toen ze/naar buiten kwamen,//
vroeg een journalist/wie van hen/Irak of Afghanistan/had/genoemd.//
Bush antwoordde/dat Clinton/Irak of Afghanistan/had/genoemd./De journalist/zei//
dat hij/dat wel/had/verwacht.

15. Bush answered / that Clinton / Iraq or Afghanistan / had / mentioned.

16. (C1)
Suzanne en Emma/kwamen terug/van vakantie./Toen Emma thuiskwam,//
vroegen/haar vriendinnen/wat voor elektronica/Suzanne/had/meegenomen.//
Emma antwoordde/dat Suzanne/een i-pod of een laptop/had/meegenomen./De vriendinnen/zeiden//
dat/dat best riskant/was.

16. (C2)
Suzanne en Emma/kwamen terug/van vakantie./Toen ze thuiskwamen,//
vroegen/hun vriendinnen/wie van hen/een i-pod of een laptop/had/meegenomen.//
Emma antwoordde/dat Suzanne/een i-pod of een laptop/had/meegenomen./De vriendinnen/zeiden//
dat/dat echt/iets voor Suzanne was.

16. Emma answered / that Suzanne / an i-pod or a laptop / had / brought.

17. (C1)
Kapitein Blauwbaard/en Kapitein Haak/gingen een schip/beroven./Toen Blauwbaard/terugkwam,//
vroegen hun mannen/welke/wapens/Haak had gebruikt.//
Blauwbaard antwoordde/dat Haak/een zwaard of een pistool/had/gebruikt./De mannen/zeiden//
dat Haak/meedogenloos was.

17. (C2)
Kapitein Blauwbaard/en Kapitein Haak/gingen/een schip/beroven./Toen ze terugkwamen,//
vroegen hun mannen/wie van hen/een zwaard of een pistool/had gebruikt.//
Blauwbaard antwoordde/dat Haak/een zwaard of een pistool/had/gebruikt./De mannen/zeiden//
dat Haak meedogenloos was.

17. Blauwbaard answered / that Hook / a sword or a gun / had / used.

18. (C1)
Lebbis/en Jansen gingen/naar een kledinginzameling/voor/arme kinderen./Toen Lebbis wegging,//
vroeg de organisator/wat voor kledingstukken/Jansen/had ingeleverd.//
Lebbis antwoordde/dat Jansen/een jas of een trui/had/ingeleverd./De organisator/zei//
dat hij/daar/blij mee was.

18. (C2)
Lebbis en Jansen/gingen naar/een kledinginzameling/voor arme kinderen./Toen ze weggingen,//
vroeg/de organisator/wie van hen/een jas of een trui/had ingeleverd.//
Lebbis antwoordde/dat Jansen/een jas of een trui/had/ingeleverd./De organisator/zei//
dat hij/verrast was.

18. Lebbis answered / that Jansen / a coat or a sweater / had / donated.

19. (C1)
Jip en Janneke/waren samen/naar de snoepwinkel/geweest./Toen Janneke/thuiskwam,//
vroeg moeder/wat voor snoepjes/Jip/had gekozen.//
Janneke antwoordde/dat Jip/perziken of banaantjes/had/gekozen./Moeder/zei//
dat dat/een goede keuze/was.

19. (C2)
Jip en Janneke/waren samen/naar de snoepwinkel geweest./Toen ze/naar buiten/kwamen,//
vroeg moeder/wie van hen/perziken of banaantjes/had gekozen.//
Janneke antwoordde/dat Jip/perziken of banaantjes/had/gekozen./Moeder/zei//
dat Jip/een zoetekauw/was.

19. Janneke answered / that Jip / peaches or bananas / had / picked.

20. (C1)
Judith en Karin/gingen/naar de MTV Music Awards./Toen/Karin terug was,//
vroegen haar klasgenoten/welke artiesten/Judith/had/aangeraakt.//
Karin antwoordde/dat Judith/Madonna of Britney Spears/had/aangeraakt./De klasgenoten/zeiden//
dat Judith/ook alles/durfde.

20. (C2)
Judith en Karin/gingen/naar de MTV Music Awards./Toen ze terug/waren,//
vroegen hun klasgenoten/wie van hen/Madonna of Britney Spears/had/aangeraakt.//
Karin antwoordde/dat Judith/Madonna of Britney Spears/had/aangeraakt./De klasgenoten/zeiden//
dat Judith/ook alles durfde.

20. Karin answered / that Judith / Madonna or Britney Spears / had / touched.

21. (C1)
Balkenende en Bos/gingen een middagje/met burgers praten./Toen Balkenende terugkwam,//
vroeg/Marijnissen/welke minderheden/Bos/had/gesproken.//
Balkenende antwoordde/dat Bos/een buitenlander of een dakloze/had/gesproken.Marijnissen/zei//
dat hij/dat dapper vond.

21. (C2)
Balkenende en Bos/gingen/een middagje/met burgers praten./Toen ze terugkwamen,//
vroeg/Marijnissen/wie van hen/een buitenlander of een dakloze/had gesproken.//
Balkenende antwoordde/dat Bos/een buitenlander of een dakloze/had/gesproken./Marijnissen/zei//
dat hij/dat niet/had/verwacht.

21. Balkenende answered / that Bos / a foreigner or a homeless person / had spoken to.

22. (C1)
Gullit en Rijkaard/gingen/naar de dierentuin./Toen/Rijkaard terugkwam,//
vroeg Van Basten/welke dieren/Gullit/had/gefotografeerd.//
Rijkaard antwoordde/dat Gullit/een leeuw of een olifant/had/gefotografeerd./Van Basten/zei//
dat hij/een hekel had/aan dierentuinen.

22. (C2)
Gullit en Rijkaard/gingen/naar de dierentuin./Toen ze terugkwamen,//
vroeg/Van Basten/wie van hen/een leeuw of een olifant/had/gefotografeerd.//
Rijkaard antwoordde/dat Gullit/een leeuw of een olifant/had/gefotografeerd./Van Basten/zei//
dat hij/Gullit daar wel voor/aanzag.

22. Rijkaard answered / that Gullit / a lion or an elephant / had / photographed.

23. (C1)
Lanceloet en Walewein/gingen naar/een feestmaal/van Koning Arthur./Toen Lanceloet wegging,//
vroeg/de koning/wat voor eten/Walewein/had gegeten.//
Lanceloet antwoordde/dat Walewein/kip of rund/had/gegeten./De koning/zei//
dat Walewein/dat/beter/niet had kunnen doen.

23. (C2)
Lanceloet en Walewein/gingen naar/een feestmaal/van Koning Arthur./Na het diner,//
vroeg/de koning/wie van hen/kip of rund/had gegeten.//
Lanceloet antwoordde/dat Walewein/kip of rund/had/gegeten./De koning/zei//
dat Walewein/dat beter/niet had kunnen doen.

23. Lanceloet answered / that Walewein / chicken or beef / had / eaten.

24. (C1)
Ingmar en Stefan/gingen/naar een spelletjesavond./Toen Ingmar terugkwam,//
vroegen/zijn ouders/welke spellen/Stefan/had/gespeeld.//
Ingmar antwoordde/dat Stefan/Risk of Monopoly/had/gespeeld./De ouders/zeiden//
dat ze/dat best ouderwets/vonden.

24. (C2)
Ingmar en Stefan/gingen/naar een spelletjesavond./Toen ze/terugkwamen,//
vroegen hun ouders/wie van hen/Risk of Monopoly/had/gespeeld.//
Ingmar antwoordde/dat Stefan/Risk of Monopoly/had/gespeeld./De ouders/zeiden//
dat ze/dat al verwacht/hadden.

24. Ingmar answered / that Stefan / Risk or Monopoly / had / played.

APPENDIX 7

Items of Experiment 9 (chapter 6)

C1 = FocRel or them
C2 = FocRel or it
C3 = non-FocRel or them
C4 = non-FocRel or it
Only the *or* conditions are reported as:
C5 (FocRel and them) = C1 with all occurrences of *or* replaced by *and*
C4 (non-FocRel and them) = C3 with all occurrences of *or* replaced by *and*

I only translate all conditions of item 1, only condition 1 for items 1-18 (as it is obvious how the other conditions are derived from this condition) and only the target sentence for items 19-36. I use the Dutch word order (and indicate the segment breaks) in the target sentences. The Dutch items contain all the segment breaks. Items 1-16 are based on the items of Experiment 5 and 7 (see Appendix 5).

1. (C1)
Karl vertelde/aan/een/vriend/dat Hugo en Olivier/gisteren/zeedieren/waren gaan//
zoeken/op het strand,/en dat/toen ze terugkwamen,/hun vader/zei/dat als/één van hen//
minstens twee zeedieren/had gevonden/op/het strand,/diegene/later mocht/opblijven.//
De vriend/zei:/"En,/wat/had/Hugo/gevonden/op het/strand?"//
Karl antwoordde:/"Hugo/had/een zeester of een mossel/gevonden/op/het/strand."//
Hugo/zei later/dat hij/ze/had gevonden/in/de/branding/bij de/vuurtoren.

1. (C1) Karl told a friend that Hugo and Oliver went searching for marine animals on the beach yesterday, and that when they returned, their father said that if one of them had found at least two marine animals on the beach, that person got to stay up late that night.
The friend said: "And, and what had Hugo found on the beach?"
Karl answered: / "Hugo / had / a starfish or a mussel / found / on / the / beach." //
Hugo / said later / that he / them / had found / in / the / surf / near the / lighthouse.

1. (C2)
Marije vertelde/aan/een/vriendin/dat Laura en Barbara/gisteren/zeedieren/waren gaan//
zoeken/op het strand,/en dat/toen ze terugkwamen,/hun moeder/zei/dat als/één van hen//
minstens twee zeedieren/had gevonden/op/het strand,/diegene/later mocht/opblijven.//
De vriendin/zei:/"En,/wat/had/Laura/gevonden/op het/strand?"//
Marije antwoordde:/"Laura/had/een zeester of een mossel/gevonden/op/het/strand."//
Laura/zei later/dat ze/'m/had gevonden/in/de/branding/bij de/vuurtoren.

1. (C2) Marije told a friend (fem.) that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found at least two marine animals on the beach, that person got to stay up late that night.
The friend (fem.) said: "And, what had Laura found on the beach?"
Marije answered: / "Laura / had / a crab or a mussel / found / on / the / beach." //
Laura / said later / that she / it / had found / in / the / surf / near the / lighthouse.

1. (C3)
Karl vertelde/aan/een/vriend/dat Hugo en Olivier/gisteren/zeedieren/waren gaan//
zoeken/op het strand,/en dat/toen ze terugkwamen,/hun vader/zei/dat als/één van hen//
een zeester of een mossel/had gevonden/op/het strand,/diegene/later mocht/opblijven.//
De vriend/zei:/"En,/wie van hen/had/een zeester of een mossel/gevonden/op het/strand?"//
Karl antwoordde:/"Hugo/had/een zeester of een mossel/gevonden/op/het/strand."//
Hugo/zei later/dat hij/ze/had gevonden/in/de/branding/bij de/vuurtoren.

1. (C3) Karl told a friend (masc.) that Hugo and Oliver went searching for marine animals on the beach yesterday, and that when they returned, their father said that if one of them had found a starfish or a mussel on the beach, that person got to stay up late that night.
The friend (masc.) said: "And, who of them had found a starfish or a mussel on the beach?"
Karl answered: / "Hugo / had / a starfish or a mussel / found / on / the / beach." //
Hugo / said later / that he / them / had found / in / the / surf / near the / lighthouse.

1. (C4)
Marije vertelde/aan/een/vriendin/dat Laura en Barbara/gisteren/zeedieren/waren gaan//
zoeken/op het strand,/en dat/toen ze terugkwamen,/hun moeder/zei/dat als/één van hen//
een zeester of een mossel/had gevonden/op/het strand,/diegene/later mocht/opblijven.//
De vriendin/zei:/"En,/wie van hen/had/een zeester of een mossel/gevonden/op het/strand?"//
Marije antwoordde:/"Laura/had/een zeester of een mossel/gevonden/op/het/strand."//
Laura/zei later/dat ze/'m/had gevonden/in/de/branding/bij de/vuurtoren.

1. (C4) Marije told a friend (fem.) that Laura and Barbara went searching for marine animals on the beach yesterday, and that when they returned, their mother said that if one of them had found a starfish or a mussel on the beach, that person got to stay up late that night.
The friend (fem.) said: "And, who of them had found a starfish or a mussel on the beach?"
Marije answered: / "Laura / had / a crab or a mussel / found / on / the / beach." //
Laura / said later / that she / it / had found / in / the / surf / near the / lighthouse.

2. (C1)
Frans/vertelde/aan/een vriend dat/Chris en Marco/gisteren/naar de snackbar/waren//
geweest,/maar dat/toen/ze terugkwamen,/hun vader/zei/dat als/één van hen//
meer dan één snack/had genomen/bij de/snackbar,/diegene/geen avondeten/zou krijgen.//
De vriend/zei:/"En,/wat/had/Marco/genomen/bij de/snackbar?"//
Frans antwoordde:/"Marco/had/een kroket of een frikadel/genomen/bij/de/snackbar."//
Marco/zei later/dat hij/ze/had genomen/uit/een/aanval/van/lekkere trek.

2. (C1) Frans told a friend that Chris and Marco went to the snack bar yesterday, but that when they returned, their father said that if one of them had taken more than one snack at the snack bar, that person would not get dinner.
The friend said: "And, what had Marco taken at the snack bar?"
Frans answered: /"Marco / had / a croquet or a meat roll / taken / at / the / snackbar. //
Marco / said later / that he / them / had taken / out of / a / sudden attack / of / the munchies.

2. (C2)
Nathalie/vertelde/aan/een vriendin dat/Hanneke en Linda/gisteren/naar de snackbar/waren//
geweest,/maar dat/toen/ze terugkwamen,/hun moeder/zei/dat als/één van hen//
meer dan één snack/had genomen/bij de/snackbar,/diegene/geen avondeten/zou krijgen.//
De vriendin/zei:/"En,/wat/had/Linda/genomen/bij de/snackbar?"//
Nathalie antwoordde:/"Linda/had/een kroket of een frikadel/genomen/bij/de/snackbar."//
Linda/zei later/dat ze/'m/had genomen/uit/een/aanval/van/lekkere trek.

2. (C3)
Frans/vertelde/aan/een vriend dat/Chris en Marco/gisteren/naar de snackbar/waren//
geweest,/maar dat/toen/ze terugkwamen,/hun vader/zei/dat als/één van hen//
een kroket of een frikadel/had genomen/bij de/snackbar,/diegene/geen avondeten/zou krijgen.//
De vriend/zei:/"En,/wie van hen/had/een kroket of een frikadel/genomen/bij de/snackbar?"//
Frans antwoordde:/"Marco/had/een kroket of een frikadel/genomen/bij/de/snackbar."//
Marco/zei later/dat hij/ze/had genomen/uit/een/aanval/van/lekkere trek.

2. (C4)
Nathalie/vertelde/aan/een vriendin dat/Hanneke en Linda/gisteren/naar de snackbar/waren//
geweest,/maar dat/toen/ze terugkwamen,/hun moeder/zei/dat als/één van hen//
een kroket of een frikadel/had genomen/bij de/snackbar,/diegene/geen avondeten/zou krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een kroket of een frikadel/genomen/bij de/snackbar?"//
Nathalie antwoordde:/"Linda/had/een kroket of een frikadel/genomen/bij/de/snackbar."//
Linda/zei later/dat ze/'m/had genomen/uit/een/aanval/van/lekkere trek.

3. (C1)
Guido/vertelde/aan een/vriend/dat Patrick en Maurice/gisteren/waren gaan/vissen/in de//
rivier,/en dat/hun vader/na afloop/zei dat/als één/van hen/minstens twee vissen//
had gevangen/in de rivier,/diegene/een/tientje/zou krijgen.//
De vriend/zei:/"En,/wat/had/Patrick/gevangen/in de/rivier?"//
Guido antwoordde:/"Patrick/had/een karper of een snoekbaars/gevangen/in/de/rivier."//
Patrick/zei later/dat hij/ze/had gevangen/met/een/hengel/zonder dobber/en zonder aas.

3. (C1) Guido told a friend that Patrick and Maurice went fishing in the river yesterday, and that their father had said afterwards that if one of them had caught at least two fish, that person would get ten bucks.
The friend said: "And, what had Patrick caught in the river?"
Guido answered: / "Patrick / had / a carp of a pike / caught / in / the / river."//
Patrick / said later / that he / them / had caught / with / a / fishing rod / without a float / and without bait.

3. (C2)
Astrid/vertelde/aan een/vriendin/dat Marloes en Francisca/gisteren/waren gaan/vissen/in de//
rivier,/en dat/hun moeder/na afloop/zei dat/als één/van hen/minstens twee vissen//
had gevangen/in de rivier,/diegene/een/tientje/zou krijgen.//
De vriendin/zei:/"En,/wat/had/Marloes/gevangen/in de/rivier?"//
Astrid antwoordde:/"Marloes/had/een karper of een snoekbaars/gevangen/in/de/rivier."//
Marloes/zei later/dat ze/'m/had gevangen/met/een/hengel/zonder dobber/en zonder aas.

3. (C3)
Guido/vertelde/aan een/vriend/dat Patrick en Maurice/gisteren/waren gaan/vissen/in de//
rivier,/en dat/hun vader/na afloop/zei dat/als één/van hen/een karper of een snoekbaars//
had gevangen/in de rivier,/diegene/een/tientje/zou krijgen.//
De vriend/zei:/"En,/wie van hen/had/een karper of een snoekbaars/gevangen/in de/rivier?"//
Guido antwoordde:/"Patrick/had/een karper of een snoekbaars/gevangen/in/de/rivier."//
Patrick/zei later/dat hij/ze/had gevangen/met/een/hengel/zonder dobber/en zonder aas.

3. (C4)
Astrid/vertelde/aan een/vriendin/dat Marloes en Francisca/gisteren/waren gaan/vissen/in de//
rivier,/en dat/hun moeder/na afloop/zei dat/als één/van hen/een karper of een snoekbaars //
had gevangen/in de rivier,/diegene/een/tientje/zou krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een karper of een snoekbaars/gevangen/in de/rivier?"//
Astrid antwoordde:/"Marloes/had/een karper of een snoekbaars/gevangen/in/de/rivier."//
Marloes/zei later/dat ze/'m/had gevangen/met/een/hengel/zonder dobber/en zonder aas.

4. (C1)
Ray/vertelde/aan een vriend/dat Dirk en Friso/vorige week/kunst/waren gaan/maken//
op een/kunstcursus,/en dat/de docent/aan/het einde/zei dat/degene die//
meerdere kunstwerken/had gemaakt/op/de cursus,/extra/materiaalkosten/moest betalen.//
De vriend/zei:/"En,/wat/had/Friso/gemaakt/op de/cursus?"//
Ray antwoordde:/"Friso/had/een ets of een collage/gemaakt/op/de/cursus."//
Friso/zei later/dat hij/ze/had gemaakt/voor/een/kennis/die ziek/was.

4. (C1) Ray told a friend that Dirk and Friso had been making art last week at an art course, and that the teacher at the end said that the one who had made more than one work of art had to pay extra for the materials.
The friend said: "And, what had Friso made at the course?"
Ray answered: / "Friso / had / an etching or a montage / made / at / the / course." //
Friso / said later / that he / them / had made / for / an / acquaintance / who sick / was.

4. (C2)
Fleur/vertelde/aan een vriendin/dat Marijne en Sophie/vorige week/kunst/waren gaan/maken//
op een/kunstcursus,/en dat/de docente/aan/het einde/zei dat/degene die//
meerdere kunstwerken/had gemaakt/op/de cursus,/extra/materiaalkosten/moest betalen.//
De vriendin/zei:/"En,/wat/had/Sophie/gemaakt/op de/cursus?"//
Fleur antwoordde:/"Sophie/had/een ets of een collage/gemaakt/op/de/cursus."//
Sophie/zei later/dat ze/'m/had gemaakt/voor/een/kennis/die ziek/was.

4. (C3)
Ray/vertelde/aan een vriend/dat Dirk en Friso/vorige week/kunst/waren gaan/maken//
op een/kunstcursus,/en dat/de docent/aan/het einde/zei dat/degene die//
een ets of een collage/had gemaakt/op/de cursus,/extra/materiaalkosten/moest betalen.//
De vriend/zei:/"En,/wie van hen/had/een ets of een collage/gemaakt/op de/cursus?"//
Ray antwoordde:/"Friso/had/een ets of een collage/gemaakt/op/de/cursus."//
Friso/zei later/dat hij/ze/had gemaakt/voor/een/kennis/die ziek/was.

4. (C4)
Fleur/vertelde/aan een vriendin/dat Marijne en Sophie/vorige week/kunst/waren gaan/maken//
op een/kunstcursus,/en dat/de docente/aan/het einde/zei dat/degene die//
een ets of een collage/had gemaakt/op/de cursus,/extra/materiaalkosten/moest betalen.//
De vriendin/zei:/"En,/wie van hen/had/een ets of een collage/gemaakt/op de/cursus?"//
Fleur antwoordde:/"Sophie/had/een ets of een collage/gemaakt/op/de/cursus."//
Sophie/zei later/dat ze/'m/had gemaakt/voor/een/kennis/die ziek/was.

5. (C1)
Nick vertelde/aan/een vriend/dat Jaap en Lennart/gisteren/in de werkplaats/elektronische//
apparaten/waren/gaan repareren,/en dat/aan het/einde van de dag/de baas/zei dat/degene die//
minstens twee apparaten/had gerepareerd/in de/werkplaats,/de dag erna/vrij/zou krijgen.//
De vriend/zei:/"En,/wat/had/Lennart/gerepareerd/in de/werkplaats?"//
Nick antwoordde:/"Lennart/had/een TV of een radio/gerepareerd/in/de/werkplaats."//
Lennart/zei later/dat hij/ze/had gerepareerd/met/een/stukje/tape/en een tangetje.

5. (C1) Nick told a friend that Jaap and Lennart were fixing electronic appliances in the workshop yesterday, and that at the end of the day the boss said that the one who had fixed at least two appliances in the workshop, would get the next day off.
The friend said: "And, what had Lennart fixed in the workshop?"
Nick answered: / "Lennart / had / a TV or a radio / fixed / in / the / workshop." //
Lennart / said later / that he / them / had fixed / with / a / piece / of tape / and pliers.

5. (C2)
Ans vertelde/aan/een vriendin/dat Esther en Sandra/gisteren/in de werkplaats/elektronische//
apparaten/waren/gaan repareren,/en dat/aan het/einde van de dag/de bazin/zei dat/degene die//
minstens twee apparaten/had gerepareerd/in de/werkplaats,/de dag erna/vrij/zou krijgen.//
De vriendin/zei:/"En,/wat/had/Sandra/gerepareerd/in de/werkplaats?"//
Ans antwoordde:/"Sandra/had/een TV of een radio/gerepareerd/in/de/werkplaats."//
Sandra/zei later/dat ze/'m/had gerepareerd/met/een/stukje/tape/en een tangetje.

5. (C3)
Nick vertelde/aan/een vriend/dat Jaap en Lennart/gisteren/in de werkplaats/elektronische//
apparaten/waren/gaan repareren,/en dat/aan het/einde van de dag/de baas/zei dat/degene die//
een TV of een radio/had gerepareerd/in de/werkplaats,/de dag erna/vrij/zou krijgen.//
De vriend/zei:/"En,/wie van hen/had/een TV of een radio/gerepareerd/in de/werkplaats?"//
Nick antwoordde:/"Lennart/had/een TV of een radio/gerepareerd/in/de/werkplaats."//
Lennart/zei later/dat hij/ze/had gerepareerd/met/een/stukje/tape/en een tangetje.

5. (C4)
Ans vertelde/aan/een vriendin/dat Esther en Sandra/gisteren/in de werkplaats/elektronische//
apparaten/waren/gaan repareren,/en dat/aan het/einde van de dag/de bazin/zei dat/degene die//
een TV of een radio/had gerepareerd/in de/werkplaats,/de dag erna/vrij/zou krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een TV of een radio/gerepareerd/in de/werkplaats?"//
Ans antwoordde:/"Sandra/had/een TV of een radio/gerepareerd/in/de/werkplaats."//
Sandra/zei later/dat ze/'m/had gerepareerd/met/een/stukje/tape/en een tangetje.

6. (C1)
Dennis/vertelde/aan een vriend/dat Frank en Robert/gisteren naar/een kledinginzameling/voor//
arme kinderen/waren/geweest,/en dat/aan het einde/de organisator/zei/dat als/één van hen//
meerdere kledingstukken/had geschonken/bij de/inzameling,/diegene/een aandenken/zou krijgen.//
De vriend/zei:/"En,/wat/had/Robert/geschonken/bij de/inzameling?"//
Dennis antwoordde:/"Robert/had/een trui of een winterjas/geschonken/bij/de/inzameling."//
Robert/zei later/dat hij/ze/had geschonken/uit/een/gevoel/van intens/medelijden.

6. (C1) Dennis told a friend that Frank and Robert had been to an event where they collected second hand clothes for poor children, and that at the end the organizer said that if one of them had donated more than one piece of clothing at the collection, that person would get a souvernir.
The friend said: "And, what had Robert donated at the collection?"
Dennis answered: / "Robert / had / a sweater or a winter coat / donated / at / the / collection." //
Robert / said later / that he / them / had donated / out of / a / feeling / of intense / pity.

6. (C2)
Anouk/vertelde/aan een vriendin/dat Saskia en Lisa/gisteren naar/een kledinginzameling/voor//
arme kinderen/waren/geweest,/en dat/aan het einde/de organisatrice/zei/dat als/één van hen//
meerdere kledingstukken/had geschonken/bij de/inzameling,/diegene/een aandenken/zou krijgen.//
De vriendin/zei:/"En,/wat/had/Lisa/geschonken/bij de/inzameling?"//
Anouk antwoordde:/"Lisa/had/een trui of een winterjas/geschonken/bij/de/inzameling."//
Lisa/zei later/dat ze/'m/had geschonken/uit/een/gevoel/van intens/medelijden.

6. (C3)
Dennis/vertelde/aan een vriend/dat Frank en Robert/gisteren naar/een kledinginzameling/voor//
arme kinderen/waren/geweest,/en dat/aan het einde/de organisator/zei/dat als/één van hen//
een trui of een winterjas/had geschonken/bij de/inzameling,/diegene/een aandenken/zou krijgen.//
De vriend/zei:/"En,/wie van hen/had/een trui of een winterjas/geschonken/bij de/inzameling?"//
Dennis antwoordde:/"Robert/had/een trui of een winterjas/geschonken/bij/de/inzameling."//
Robert/zei later/dat hij/ze/had geschonken/uit/een/gevoel/van intens/medelijden.

6. (C4)
Anouk/vertelde/aan een vriendin/dat Saskia en Lisa/gisteren naar/een kledinginzameling/voor//
arme kinderen/waren/geweest,/en dat/aan het einde/de organisatrice/zei/dat als/één van hen//
een trui of een winterjas/had geschonken/bij de/inzameling,/diegene/een aandenken/zou krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een trui of een winterjas/geschonken/bij de/inzameling?"//
Anouk antwoordde:/"Lisa/had/een trui of een winterjas/geschonken/bij/de/inzameling."//
Lisa/zei later/dat ze/'m/had geschonken/uit/een/gevoel/van intens/medelijden.

7. (C1)
Johan/vertelde/aan/een vriend/dat Bastiaan en Gerard/gisteren/naar/de/Ikea/waren gegaan,//
en dat/er/een actie was/dat als/je/meerdere meubels/had/gekocht/bij de Ikea,/je//
kans maakte/op/een prijs.//
De vriend/zei:/"En,/wat/had/Gerard/gekocht/bij de/Ikea?"//
Johan antwoordde:/"Gerard/had/een bank of een eettafel/gekocht/bij/de/Ikea."//
Gerard/zei later/dat hij/ze/had gekocht/voor/de/woning/in aanbouw/in Almere.

7. (C1) Johan told a friend that Bastiaan and Gerard went to Ikea yesterday, and that there was a promotion that if you had bought more than one piece of furniture at Ikea, you competed for a prize.
The friend said: "And, what had Gerard bought at Ikea?"
Johan answered: / "Gerard / had / a couch or a dining table / bought / at / the / Ikea." // Gerard / said later / that he / them / had bought / for / the / house / under construction / in Almere.

7. (C2)
Sanne/vertelde/aan/een vriendin/dat Jenny en Marjan/gisteren/naar/de/Ikea/waren gegaan,//
en dat/er/een actie was/dat als/je/meerdere meubels/had/gekocht/bij de Ikea,/je//
kans maakte/op/een prijs.//
De vriendin/zei:/"En,/wat/had/Jenny/gekocht/bij de/Ikea?"//
Sanne antwoordde:/"Jenny/had/een bank of een eettafel/gekocht/bij/de/Ikea."//
Jenny/zei later/dat ze/'m/had gekocht/voor/de/woning/in aanbouw/in Almere.

7. (C3)
Johan/vertelde/aan/een vriend/dat Bastiaan en Gerard/gisteren/naar/de/Ikea/waren gegaan,//
en dat/er/een actie was/dat als/je/een bank of een eettafel/had/gekocht/bij de Ikea,/je//
kans maakte/op/een prijs.//
De vriend/zei:/"En,/wie van hen/had/een bank of een eettafel/gekocht/bij de/Ikea?"//
Johan antwoordde:/"Gerard/had/een bank of een eettafel/gekocht/bij/de/Ikea."//
Gerard/zei later/dat hij/ze/had gekocht/voor/de/woning/in aanbouw/in Almere.

7. (C4)
Sanne/vertelde/aan/een vriendin/dat Jenny en Marjan/gisteren/naar/de/Ikea/waren gegaan,//
en dat/er/een actie was/dat als/je/een bank of een eettafel/had/gekocht/bij de Ikea,/je//
kans maakte/op/een prijs.//
De vriendin/zei:/"En,/wie van hen/had/een bank of een eettafel/gekocht/bij de/Ikea?"//
Sanne antwoordde:/"Jenny/had/een bank of een eettafel/gekocht/bij/de/Ikea."//
Jenny/zei later/dat ze/'m/had gekocht/voor/de/woning/in aanbouw/in Almere.

8. (C1)
Tim/vertelde aan/een vriend/dat/Mark en Herman/gisteren/hadden meegedaan/aan een//
muziekworkshop,/en dat/de leraar/aan het/einde/zei/dat/degene die//
meerdere instrumenten/had/geprobeerd/bij de/workshop,/een applaus/kreeg.//
De vriend/zei:/"En,/wat/had/Herman/geprobeerd/bij de/workshop?"//
Tim antwoordde:/"Herman/had/een saxofoon of een trompet/geprobeerd/bij/de/workshop."//
Herman/zei later/dat hij/ze/had geprobeerd/aan/het/begin/van de/workshop.

8. (C1) Tim told a friend that Mark and Herman participated in a music workshop yesterday, and that the teacher said at the end that the person who had tried more than one musical instrument at the workshop, would be rewarded with an applause.
The friend said: "And, what had Herman tried at the workshop?"
Tim answered: / "Herman / had / a saxophone or a trumpet / tried / at / the / workshop." //
Herman / said later / that he / them / had tried / at / the / beginning / of the / workshop.

8. (C2)
Carolien/vertelde aan/een vriendin/dat/Miranda en Yvette/gisteren/hadden meegedaan/aan een//
muziekworkshop,/en dat/de lerares/aan het/einde/zei/dat/degene die//
meerdere instrumenten/had/geprobeerd/bij de/workshop,/een applaus/kreeg.//
De vriendin/zei:/"En,/wat/had/Yvette/geprobeerd/bij de/workshop?"//
Carolien antwoordde:/"Yvette/had/een saxofoon of een trompet/geprobeerd/bij/de/workshop."//
Yvette/zei later/dat ze/'m/had geprobeerd/aan/het/begin/van de/workshop.

8. (C3)
Tim/vertelde aan/een vriend/dat/Mark en Herman/gisteren/hadden meegedaan/aan een//
muziekworkshop,/en dat/de leraar/aan het/einde/zei/dat/degene die//
een saxofoon of een trompet/had/geprobeerd/bij de/workshop,/een applaus/kreeg.//
De vriend/zei:/"En,/wie van hen/had/een saxofoon of een trompet/geprobeerd/bij de/workshop?"//
Tim antwoordde:/"Herman/had/een saxofoon of een trompet/geprobeerd/bij/de/workshop."//
Herman/zei later/dat hij/ze/had geprobeerd/aan/het/begin/van de/workshop.

8. (C4)
Carolien/vertelde aan/een vriendin/dat/Miranda en Yvette/gisteren/hadden meegedaan/aan een//
muziekworkshop,/en dat/de lerares/aan het/einde/zei/dat/degene die//
een saxofoon of een trompet/had/geprobeerd/bij de/workshop,/een applaus/kreeg.//
De vriendin/zei:/"En,/wie van hen/had/een saxofoon of een trompet/geprobeerd/bij de/workshop?"//
Carolien antwoordde:/"Yvette/had/een saxofoon of een trompet/geprobeerd/bij/de/workshop."//
Yvette/zei later/dat ze/'m/had geprobeerd/aan/het/begin/van de/workshop.

9. (C1)
Martin vertelde/aan een vriend/dat Steven en Roy/gisteren/op een cursus/zelf/kleding//
waren/gaan maken,/en dat/de leraar/aan het einde/zei dat/degene/die//
meerdere kledingstukken/had gemaakt/op de/cursus,/een extra hoge/beoordeling/zou/krijgen.//
De vriend/zei:/"En,/wat/had/Steven/gemaakt/op de/cursus?"//
Martin antwoordde:/"Steven/had/een pantalon of een blouse/gemaakt/op/de/cursus."//
Steven/zei later/dat hij/ze/had gemaakt/van/een/stofje/met bloemetjes/en hartjes.

9. (C1) Martin told a friend that Steven and Roy had been making clothes at a handicraft course yesterday and that the teacher had said at the end that the one who had made more than one piece of clothing at the course, would get a high evaluation.
The friend said: "And, what had Steven made at the course?"
Martin answered: / "Steve / had / pants or a blouse / made / at / the / course." //
Steve / said later / that he / them / had made / out of / a / fabric / with flowers / and harts.

9. (C2)
Marie vertelde/aan een vriendin/dat Wendy en Ingrid/gisteren/op een cursus/zelf/kleding//
waren/gaan maken,/en dat/de lerares/aan het einde/zei dat/degene/die//
meerdere kledingstukken/had gemaakt/op de/cursus,/een extra hoge/beoordeling/zou/krijgen.//
De vriendin/zei:/"En,/wat/had/Wendy/gemaakt/op de/cursus?"//
Marie antwoordde:/"Wendy/had/een pantalon of een blouse/gemaakt/op/de/cursus."//
Wendy/zei later/dat ze/'m/had gemaakt/van/een/stofje/met bloemetjes/en hartjes.

9. (C3)
Martin vertelde/aan een vriend/dat Steven en Roy/gisteren/op een cursus/zelf/kleding//
waren/gaan maken,/en dat/de leraar/aan het einde/zei dat/degene/die//
een pantalon of een blouse/had gemaakt/op de/cursus,/een extra hoge/beoordeling/zou/krijgen.//
De vriend/zei:/"En,/wie van hen/had/een pantalon of een blouse/gemaakt/op de/cursus?"//
Martin antwoordde:/"Steven/had/een pantalon of een blouse/gemaakt/op/de/cursus."//
Steven/zei later/dat hij/ze/had gemaakt/van/een/stofje/met bloemetjes/en hartjes.

9. (C4)
Marie vertelde/aan een vriendin/dat Wendy en Ingrid/gisteren/op een cursus/zelf/kleding//
waren/gaan maken,/en dat/de lerares/aan het einde/zei dat/degene/die//
een pantalon of een blouse/had gemaakt/op de/cursus,/een extra hoge/beoordeling/zou/krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een pantalon of een blouse/gemaakt/op de/cursus?"//
Marie antwoordde:/"Wendy/had/een pantalon of een blouse/gemaakt/op/de/cursus."//
Wendy/zei later/dat ze/'m/had gemaakt/van/een/stofje/met bloemetjes/en hartjes.

10. (C1)
Peter/vertelde/aan een/vriend dat/Richard en Wilco/gisteren/naar/een beurs/voor//
computeronderdelen/waren gegaan,/en dat/als je/meerdere artikelen/had/gekocht//
op/de beurs,/je/een kortingsbon/voor/een computerwinkel/kreeg.//
De vriend/zei:/"En,/wat/had/Richard/gekocht/op de/beurs?"//
Peter antwoordde:/"Richard/had/een toetsenbord of een muis/gekocht/op/de/beurs."//
Richard/zei later/dat hij/ze/had gekocht/om/de/klachten/van RSI/tegen te gaan.

10. (C1) Peter told a friend that Richard and Wilco had been to a fair for computer parts yesterday, and that if one bought more than one computer part at the fair, one got a discount cheque for a computer store.
The friend said: "And, what had Richard bought at the fair?"
Peter answered: / "Richard / had / a keyboard or a mouse / bought / at / the / fair." //
Richard / said later / that he / them / had bought / to / the / symptoms / of RSI / get rid of.

10. (C2)
Yvonne/vertelde/aan een/vriendin dat/Rianne en Nicole/gisteren/naar/een beurs/voor//
computeronderdelen/waren gegaan,/en dat/als je/meerdere artikelen/had/gekocht//
op/de beurs,/je/een kortingsbon/voor/een computerwinkel/kreeg.//
De vriendin/zei:/"En,/wat/had/Rianne/gekocht/op de/beurs?"//
Yvonne antwoordde:/"Rianne/had/een toetsenbord of een muis/gekocht/op/de/beurs."//
Rianne/zei later/dat ze/'m/had gekocht/om/de/klachten/van RSI/tegen te gaan.

10. (C3)
Peter/vertelde/aan een/vriend dat/Richard en Wilco/gisteren/naar/een beurs/voor//
computeronderdelen/waren gegaan,/en dat/als je/een toetsenbord of een muis/had/gekocht//
op/de beurs,/je/een kortingsbon/voor/een computerwinkel/kreeg.//
De vriend/zei:/"En,/wie van hen/had/een toetsenbord of een muis/gekocht/op de/beurs?"//
Peter antwoordde:/"Richard/had/een toetsenbord of een muis/gekocht/op/de/beurs."//
Richard/zei later/dat hij/ze/had gekocht/om/de/klachten/van RSI/tegen te gaan.

10. (C4)
Yvonne/vertelde/aan een/vriendin dat/Rianne en Nicole/gisteren/naar/een beurs/voor//
computeronderdelen/waren gegaan,/en dat/als je/een toetsenbord of een muis/had/gekocht//
op/de beurs,/je/een kortingsbon/voor/een computerwinkel/kreeg.//
De vriendin/zei:/"En,/wie van hen/had/een toetsenbord of een muis/gekocht/op de/beurs?"//
Yvonne antwoordde:/"Rianne/had/een toetsenbord of een muis/gekocht/op/de/beurs."//
Rianne/zei later/dat ze/'m/had gekocht/om/de/klachten/van RSI/tegen te gaan.

11. (C1)
Emiel/vertelde/aan een/vriend/dat Karel en Bart/op vakantie/enge/beesten/waren gaan//
vangen/in de jungle,/en dat/toen ze/terugkwamen,/de reisleider/zei/dat degene die//
meer dan één eng beest/had gevangen/in de jungle,/een/beloning/zou krijgen.//
De vriend/zei:/"En,/wat/had/Karel/gevangen/in de/jungle?"//
Emiel antwoordde:/"Karel/had/een spin of een schorpioen/gevangen/in/de/jungle."//
Karel/zei later/dat hij/ze/had gevangen/bij/een/beekje/tussen de/stenen.

11. (C1) Emiel told a friend that Karel and Bart went out to catch creepy animals in the jungle during their vacation, and that when they returned, the travel guide said that the one who had caught more than one creepy animal in the jungle, would get a reward.
The friend said: "And, what had Karel caught in the jungle?"
Emiel answered: /"Karel / had / a spider or a scorpion / caught / in / the / jungle."
Karel / said later / that he / them / had caught / near / a stream / between the / rocks.

11. (C2)
Anita/vertelde/aan een/vriendin/dat Julie en Karin/op vakantie/enge/beesten/waren gaan//
vangen/in de jungle,/en dat/toen ze/terugkwamen,/de reisleidster/zei/dat degene die//
meer dan één eng beest/had gevangen/in de jungle,/een/beloning/zou krijgen.//
De vriendin/zei:/"En,/wat/had/Julie/gevangen/in de/jungle?"//
Anita antwoordde:/"Julie/had/een spin of een schorpioen/gevangen/in/de/jungle."//
Julie/zei later/dat ze/'m/had gevangen/bij/een/beekje/tussen de/stenen.

11. (C3)
Emiel/vertelde/aan een/vriend/dat Karel en Bart/op vakantie/enge/beesten/waren gaan//
vangen/in de jungle,/en dat/toen ze/terugkwamen,/de reisleider/zei/dat degene die//
een spin of een schorpioen/had gevangen/in de jungle,/een/beloning/zou krijgen.//
De vriend/zei:/"En,/wie van hen/had/een spin of een schorpioen/gevangen/in de/jungle?"//
Emiel antwoordde:/"Karel/had/een spin of een schorpioen/gevangen/in/de/jungle."//
Karel/zei later/dat hij/ze/had gevangen/bij/een/beekje/tussen de/stenen.

11. (C4)
Anita/vertelde/aan een/vriendin/dat Julie en Karin/op vakantie/enge/beesten/waren gaan//
vangen/in de jungle,/en dat/toen ze/terugkwamen,/de reisleidster/zei/dat degene die//
een spin of een schorpioen/had gevangen/in de jungle,/een/beloning/zou krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een spin of een schorpioen/gevangen/in de/jungle?"//
Anita antwoordde:/"Julie/had/een spin of een schorpioen/gevangen/in/de/jungle."//
Julie/zei later/dat ze/'m/had gevangen/bij/een/beekje/tussen de/stenen.

12. (C1)
Kevin/vertelde/aan een vriend/dat/Martijn en Paul/gisteren/planten/waren gaan/kopen//
bij het/tuincentrum,/en dat/er/een actie was/dat als je/meerdere planten/had gekocht//
bij/het tuincentrum,/je korting/kreeg/bij de bouwmarkt/ernaast.//
De vriend/zei:/"En,/wat/had/Martijn/gekocht/bij het/tuincentrum?"//
Kevin antwoordde:/"Martijn/had/een cactus of een vetplant/gekocht/bij/het/tuincentrum."//
Martijn/zei later/dat hij/ze/had gekocht/voor/het/balkon/aan de straatkant/van het huis.

12. (C1) Kevin told a friend that Martijn and Paul went to buy plants yesterday at the garden center, and that there was a promotion that if you bought more than one plant at the garden center, you'd get a discount at the hardware store next door.
The friend said: "And, what had Martijn bought at the garden center?"
Kevin answered: / "Martijn / had / a cactus or a succulent / bought / at / the / garden center.// Martijn / said later / that he / them / had bought / for / the / balcony / on the street side / of the house.

12. (C2)
Claudia/vertelde/aan een vriendin/dat/Paula en Martine/gisteren/planten/waren gaan/kopen//
bij het/tuincentrum,/en dat/er/een actie was/dat als je/meerdere planten/had gekocht//
bij/het tuincentrum,/je korting/kreeg/bij de bouwmarkt/ernaast.//
De vriendin/zei:/"En,/wat/had/Paula/gekocht/bij het/tuincentrum?"//
Claudia antwoordde:/"Paula/had/een cactus of een vetplant/gekocht/bij/het/tuincentrum."//
Paula/zei later/dat ze/'m/had gekocht/voor/het/balkon/aan de straatkant/van het huis.

12. (C3)
Kevin/vertelde/aan een vriend/dat/Martijn en Paul/gisteren/planten/waren gaan/kopen//
bij het/tuincentrum,/en dat/er/een actie was/dat als je/een cactus of een vetplant/had gekocht//
bij/het tuincentrum,/je korting/kreeg/bij de bouwmarkt/ernaast.//
De vriend/zei:/"En,/wie van hen/had/een cactus of een vetplant/gekocht/bij het/tuincentrum?"//
Kevin antwoordde:/"Martijn/had/een cactus of een vetplant/gekocht/bij/het/tuincentrum."//
Martijn/zei later/dat hij/ze/had gekocht/voor/het/balkon/aan de straatkant/van het huis.

12. (C4)
Claudia/vertelde/aan een vriendin/dat/Paula en Martine/gisteren/planten/waren gaan/kopen//
bij het/tuincentrum,/en dat/er/een actie was/dat als je/een cactus of een vetplant/had gekocht//
bij/het tuincentrum,/je korting/kreeg/bij de bouwmarkt/ernaast.//
De vriendin/zei:/"En,/wie van hen/had/een cactus of een vetplant/gekocht/bij het/tuincentrum?"//
Claudia antwoordde:/"Paula/had/een cactus of een vetplant/gekocht/bij/het/tuincentrum."//
Paula/zei later/dat ze/'m/had gekocht/voor/het/balkon/aan de straatkant/van het huis.

13. (C1)
Stefan/vertelde/aan een vriend/dat Floris en Rob/gisteren/naar/de/kermis/waren geweest,//
en dat/bij de/uitgang/bleek/dat degene die/meer dan één prijs/had/gewonnen//
op/de kermis,/gratis/op de/foto/mocht.//
De vriend/zei:/"En,/wat/had/Floris/gewonnen/op de/kermis?"//
Stefan antwoordde:/"Floris/had/een knuffel of een skippybal/gewonnen/op/de/kermis."//
Floris/zei later/dat hij/ze/had gewonnen/bij/het/kraampje/van het/touwtje trekken.

13. (C1)
Stefan told a friend that Floris and Rob had been to the fair yesterday, and that at the exit it turned out that if one had won more than one prize, one could get his picture taken for free.
The friend said: "Oh, and what did Floris win at the fair?"
Stefan answered: / "Floris / had / a teddy bear or a skippy ball / won / at / the / fair."
Floris / said later / that he / them / had won / at / the / stand / of the / 'touwtje trekken'.

13. (C2)
Margot/vertelde/aan een vriendin/dat Emma en José/gisteren/naar/de/kermis/waren geweest,//
en dat/bij de/uitgang/bleek/dat degene die/meer dan één prijs/had/gewonnen//
op/de kermis,/gratis/op de/foto/mocht.//
De vriendin/zei:/"En,/wat/had/Emma/gewonnen/op de/kermis?"//
Margot antwoordde:/"Emma/had/een knuffel of een skippybal/gewonnen/op/de/kermis."//
Emma/zei later/dat ze/'m/had gewonnen/bij/het/kraampje/van het/touwtje trekken.

13. (C3)
Stefan/vertelde/aan een vriend/dat Floris en Rob/gisteren/naar/de/kermis/waren geweest,//
en dat/bij de/uitgang/bleek/dat degene die/een knuffel of een skippybal/had/gewonnen//
op/de kermis,/gratis/op de/foto/mocht.//
De vriend/zei:/"En,/wie van hen/had/een knuffel of een skippybal/gewonnen/op de/kermis?"//
Stefan antwoordde:/"Floris/had/een knuffel of een skippybal/gewonnen/op/de/kermis."//
Floris/zei later/dat hij/ze/had gewonnen/bij/het/kraampje/van het/touwtje trekken.

13. (C4)
Margot/vertelde/aan een vriendin/dat Emma en José/gisteren/naar/de/kermis/waren geweest,//
en dat/bij de/uitgang/bleek/dat degene die/een knuffel of een skippybal/had/gewonnen//
op/de kermis,/gratis/op de/foto/mocht.//
De vriendin/zei:/"En,/wie van hen/had/een knuffel of een skippybal/gewonnen/op de/kermis?"//
Margot antwoordde:/"Emma/had/een knuffel of een skippybal/gewonnen/op/de/kermis."//
Emma/zei later/dat ze/'m/had gewonnen/bij/het/kraampje/van het/touwtje trekken.

14. (C1)
Ernst/vertelde/aan een/vriend/dat Pieter en Roderick/gisteren/waren gaan/jagen,//
en dat/toen ze/terugkwamen,/de/voorzitter/van de jachtclub/zei dat/degene die//
minstens twee beesten/had/geschoten/bij de jacht,/lid/mocht worden/van de jachtclub.//
De vriend/zei:/"En,/wat/had/Pieter/geschoten/bij de/jacht?"//
Ernst antwoordde:/"Pieter/had/een fazant of een konijn/geschoten/bij/de/jacht."//
Pieter/zei later/dat hij/ze/had geschoten/van/een/afstand/van minstens/100 meter.

14. (C1) Ernst told a friend that Pieter and Roderick went out hunting for game yesterday, and that when they got back, the president of the hunting club told them that the one who had shot at least two animals at the hunt, was allowed to join the hunting club.
The friend said: "And, what had Pieter shot at the hunt?"
Ernst answered: / "Pieter / had / a pheasant or a rabbit / shot / at / the / hunt." //
Pieter / said later / that he / them / had shot / from / a distance / of at least / 100 meters.

14. (C2)
Laurien/vertelde/aan een/vriendin/dat Ellen en Leonoor/gisteren/waren gaan/jagen,//
en dat/toen ze/terugkwamen,/de/voorzitster/van de jachtclub/zei dat/degene die//
minstens twee beesten/had/geschoten/bij de jacht,/lid/mocht worden/van de jachtclub.//
De vriendin/zei:/"En,/wat/had/Ellen/geschoten/bij de/jacht?"//
Laurien antwoordde:/"Ellen/had/een fazant of een konijn/geschoten/bij/de/jacht."//
Ellen/zei later/dat ze/'m/had geschoten/van/een/afstand/van minstens/100 meter.

14. (C3)
Ernst/vertelde/aan een/vriend/dat Pieter en Roderick/gisteren/waren gaan/jagen,//
en dat/toen ze/terugkwamen,/de/voorzitter/van de jachtclub/zei dat/degene die//
een fazant of een konijn/had/geschoten/bij de jacht,/lid/mocht worden/van de jachtclub.//
De vriend/zei:/"En,/wie van hen/had/een fazant of een konijn/geschoten/bij de/jacht?"//
Ernst antwoordde:/"Pieter/had/een fazant of een konijn/geschoten/bij/de/jacht."//
Pieter/zei later/dat hij/ze/had geschoten/van/een/afstand/van minstens/100 meter.

14. (C4)
Laurien/vertelde/aan een/vriendin/dat Ellen en Leonoor/gisteren/waren gaan/jagen,//
en dat/toen ze/terugkwamen,/de/voorzitster/van de jachtclub/zei dat/degene die//
een fazant of een konijn/had/geschoten/bij de jacht,/lid/mocht worden/van de jachtclub.//
De vriendin/zei:/"En,/wie van hen/had/een fazant of een konijn/geschoten/bij de/jacht?"//
Laurien antwoordde:/"Ellen/had/een fazant of een konijn/geschoten/bij/de/jacht."//
Ellen/zei later/dat ze/'m/had geschoten/van/een/afstand/van minstens/100 meter.

15. (C1)
Rik/vertelde/aan/een vriend/dat Michiel en Nico/gisteren/op een/archeologische/afgraving//
naar oude resten/waren gaan/zoeken,/en dat de leider/aan het/einde/zei dat/degene die//
meerdere resten/had gevonden/bij de afgraving,/de volgende dag/terug/mocht komen.//
De vriend/zei:/"En,/wat/had/Nico/gevonden/bij de/afgraving?"//
Rik antwoordde:/"Nico/had/een pijlpunt of een potscherf/gevonden/bij/de/afgraving."//
Nico/zei later/dat hij/ze/had gevonden/in/een/hoekje/waar nog niemand/had gekeken.

15. (C1) Rik told a friend that Michiel and Nico had been searching for ancient remnants at an archaeological dig yesterday, and that the leider of the dig at the end of the day said that the one who had found more than one remnant at the dig, was allowed to return the next day.
The friend said: "And, what had Nico found at the dig?"
Rik answered: / "Nico / had / an arrowhead or a potsherd / found / at / the / dig." //
Nico / said later / that he / them / had found / in / a / corner / where nobody / had looked.

15. (C2)
Leonie/vertelde/aan/een vriendin/dat Charlotte en Rosa/gisteren/op een/archeologische/afgraving//
naar oude resten/waren gaan/zoeken,/en dat de leidster/aan het/einde/zei dat/degene die//
meerdere resten/had gevonden/bij de afgraving,/de volgende dag/terug/mocht komen.//
De vriendin/zei:/"En,/wat/had/Rosa/gevonden/bij de/afgraving?"//
Leonie antwoordde:/"Rosa/had/een pijlpunt of een potscherf/gevonden/bij/de/afgraving."//
Rosa/zei later/dat ze/'m/had gevonden/in/een/hoekje/waar nog niemand/had gekeken.

15. (C3)
Rik/vertelde/aan/een vriend/dat Michiel en Nico/gisteren/op een/archeologische/afgraving//
naar oude resten/waren gaan/zoeken,/en dat de leider/aan het/einde/zei dat/degene die//
een pijlpunt of een potscherf/had gevonden/bij de afgraving,/de volgende dag/terug/mocht komen.//
De vriend/zei:/"En,/wie van hen/had/een pijlpunt of een potscherf/gevonden/bij de/afgraving?"//
Rik antwoordde:/"Nico/had/een pijlpunt of een potscherf/gevonden/bij/de/afgraving."//
Nico/zei later/dat hij/ze/had gevonden/in/een/hoekje/waar nog niemand/had gekeken.

15. (C4)
Leonie/vertelde/aan/een vriendin/dat Charlotte en Rosa/gisteren/op een/archeologische/afgraving//
naar oude resten/waren gaan/zoeken,/en dat de leidster/aan het/einde/zei dat/degene die//
een pijlpunt of een potscherf/had gevonden/bij de afgraving,/de volgende dag/terug/mocht komen.//
De vriendin/zei:/"En,/wie van hen/had/een pijlpunt of een potscherf/gevonden/bij de/afgraving?"//
Leonie antwoordde:/"Rosa/had/een pijlpunt of een potscherf/gevonden/bij/de/afgraving."//
Rosa/zei later/dat ze/'m/had gevonden/in/een/hoekje/waar nog niemand/had gekeken.

16. (C1)
Bas/vertelde/aan/een vriend/dat Albert en Remco/gisteren/naar een beurs/waren/geweest,/en dat//
toen ze/terugkwamen,/hun/huisgenoot/zei dat/als één/van hen/meer dan één relatiegeschenk//
had meegenomen/van de beurs,/diegene/echt sneu/was.//
De vriend/zei:/"En,/wat/had/Albert/meegenomen/van de/beurs?"//
Bas antwoordde:/"Albert/had/een pen of een sleutelhanger/meegenomen/van/de/beurs."//
Albert/zei later/dat hij/ze/had meegenomen/bij/een/kraampje/van een/reisbureau.

16. (C1) Bas told a friend that Albert and Remco had been to a convention yesterday, and that when they returned, their housemate said that if one if them had brought back more than one promotional gift from the convention, that person was just sad.
The friend said: "And, what had Albert brought back from the convention?"
Bas answered: / "Albert / had / a pen or a keycahin / brought/ from / the / convention." //
Albert / said later / that he / them / had brought back / from / a / stand / of a / travel agency.

16. (C2)
Sylvia/vertelde/aan/een vriendin/dat Dagmar en Femke/gisteren/naar een beurs/waren/geweest,/en dat//
toen ze/terugkwamen,/hun/huisgenootje/zei dat/als één/van hen/meer dan één relatiegeschenk//
had meegenomen/van de beurs,/diegene/echt sneu/was.//
De vriendin/zei:/"En,/wat/had/Dagmar/meegenomen/van de/beurs?"//
Sylvia antwoordde:/"Dagmar/had/een pen of een sleutelhanger/meegenomen/van/de/beurs."//
Dagmar/zei later/dat ze/'m/had meegenomen/bij/een/kraampje/van een/reisbureau.

16. (C3)
Bas/vertelde/aan/een vriend/dat Albert en Remco/gisteren/naar een beurs/waren/geweest,/en dat//
toen ze/terugkwamen,/hun/huisgenoot/zei dat/als één/van hen/een pen of een sleutelhanger//
had meegenomen/van de beurs,/diegene/echt sneu/was.//
De vriend/zei:/"En,/wie van hen/had/een pen of een sleutelhanger/meegenomen/van de/beurs?"//
Bas antwoordde:/"Albert/had/een pen of een sleutelhanger/meegenomen/van/de/beurs."//
Albert/zei later/dat hij/ze/had meegenomen/bij/een/kraampje/van een/reisbureau.

16. (C4)
Sylvia/vertelde/aan/een vriendin/dat Dagmar en Femke/gisteren/naar een beurs/waren/geweest,/en dat//
toen ze/terugkwamen,/hun/huisgenootje/zei dat/als één/van hen/een pen of een sleutelhanger//
had meegenomen/van de beurs,/diegene/echt sneu/was.//
De vriendin/zei:/"En,/wie van hen/had/een pen of een sleutelhanger/meegenomen/van de/beurs?"//
Sylvia antwoordde:/"Dagmar/had/een pen of een sleutelhanger/meegenomen/van/de/beurs."//
Dagmar/zei later/dat ze/'m/had meegenomen/bij/een/kraampje/van een/reisbureau.

17. (C1)
Maarten/vertelde aan/een vriend/dat/Evert en Jeroen/gisteren/waren gaan/shoppen/bij//
de/Media Markt,/en dat/toen ze terugkwamen/hun vader zei/dat/als één van hen/meerdere artikelen//
had gekocht/bij de Media Markt,/diegene/voorlopig geen/zakgeld/zou krijgen.//
De vriend/zei:/"En,/wat/had/Evert/gekocht/bij de/Media Markt?"//
Maarten antwoordde:/"Evert/had/een CD of een DVD/gekocht/bij/de/Media Markt."//
Evert/zei later/dat hij/ze/had gekocht/voor/een/appel/en een/ei.

17. (C1)
Maarten told a friend that Evert and Jeroen went shopping at the electronics store Media Markt yesterday, and that when they returned their father said that if one of them had bought more than one item at Media Markt, that person would not get any allowance anytime soon.
The friend said: "And, what had Evert bought at the Media Markt?"
Maarten answered: / "Evert had / a CD or a DVD / bought / at / the / Media Markt." //
Evert / said later / that he / them / had bought / for / almost nothing.

17. (C2)
Janna/vertelde aan een/vriendin/dat/Judith en Sara/gisteren/waren gaan/shoppen/bij//
de/Media Markt,/en dat/toen ze terugkwamen/hun moeder zei/dat/als één van hen/meerdere artikelen//
had gekocht/bij de Media Markt,/diegene/voorlopig geen/zakgeld/zou krijgen.//
De vriendin/zei:/"En,/wat/had/Judith/gekocht/bij de/Media Markt?"//
Janna antwoordde:/"Judith/had/een CD of een DVD/gekocht/bij/de/Media Markt."//
Judith/zei later/dat ze/'m/had gekocht/voor/een/appel/en een/ei.

17. (C3)
Maarten/vertelde aan/een vriend/dat/Evert en Jeroen/gisteren/waren gaan/shoppen/bij//
de/Media Markt,/en dat/toen ze terugkwamen/hun vader zei/dat/als één van hen/een CD of een DVD//
had gekocht/bij de Media Markt,/diegene/voorlopig geen/zakgeld/zou krijgen.//
De vriend/zei:/"En,/wie van hen/had/een CD of een DVD/gekocht/bij de/Media Markt?"//
Maarten antwoordde:/"Evert/had/een CD of een DVD/gekocht/bij/de/Media Markt."//
Evert/zei later/dat hij/ze/had gekocht/voor/een/appel/en een/ei.

17. (C4)
Janna/vertelde aan een/vriendin/dat/Judith en Sara/gisteren/waren gaan/shoppen/bij//
de/Media Markt,/en dat/toen ze terugkwamen/hun moeder zei/dat/als één van hen/een CD of een DVD//
had gekocht/bij de Media Markt,/diegene/voorlopig geen/zakgeld/zou krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een CD of een DVD/gekocht/bij de/Media Markt?"//
Janna antwoordde:/"Judith/had/een CD of een DVD/gekocht/bij/de/Media Markt."//
Judith/zei later/dat ze/'m/had gekocht/voor/een/appel/en een/ei.

18. (C1)
Mitchel/vertelde/aan een vriend/dat Sjonnie en Barrie/waren betrapt/op het/stelen//
van snoep/bij de/supermarkt,/en dat/de supermarkteigenaar/zei/dat als/één van hen//
meer dan één reep/had/gestolen,/diegene/naar/het/politiebureau/moest.//
De vriend/zei:/"En,/wat/had/Barrie/gestolen/bij de/supermarkt?"//
Mitchel antwoordde:/"Barrie/had/een Snickers of een Bounty/gestolen/bij/de/supermarkt."//
Barrie/zei later/dat hij/ze/had gestolen/voor/de/spanning/en om stoer/te doen.

18. (C1) Mitchel told a friend that Sjonnie and Barrie were caught stealing candy at the supermarket, and that the owner said that if one of them had stolen more than one candy bar, that person would be sent to the police.
The friend said: "Oh, and what had Barrie stolen at the supermarket?"
Mitchel answered: / "Barrie / had / a Snickers or a Bounty / stolen / at / the / supermarket." //
Barrie / said later / that he / them / had stolen / for / the / excitement / and to be / cool.

18. (C2)
Samantha/vertelde/aan een vriendin/dat Bianca en Shirley/waren betrapt/op het/stelen//
van snoep/bij de/supermarkt,/en dat/de supermarkteigenaresse/zei/dat als/één van hen//
meer dan één reep/had/gestolen,/diegene/naar/het/politiebureau/moest.//
De vriendin/zei:/"En,/wat/had/Shirley/gestolen/bij de/supermarkt?"//
Samantha antwoordde:/"Shirley/had/een Snickers of een Bounty/gestolen/bij/de/supermarkt."//
Shirley/zei later/dat ze/'m/had gestolen/voor/de/spanning/en om stoer/te doen.
18. (C3)
Mitchel/vertelde/aan een vriend/dat Sjonnie en Barrie/waren betrapt/op het/stelen//
van snoep/bij de/supermarkt,/en dat/de supermarkteigenaar/zei/dat als/één van hen//
een Snickers of een Bounty/had/gestolen,/diegene/naar/het/politiebureau/moest.//
De vriend/zei:/"En,/wie van hen/had/een Snickers of een Bounty/gestolen/bij de/supermarkt?"//
Mitchel antwoordde:/"Barrie/had/een Snickers of een Bounty/gestolen/bij/de/supermarkt."//
Barrie/zei later/dat hij/ze/had gestolen/voor/de/spanning/en om stoer/te doen.

18. (C4)
Samantha/vertelde/aan een vriendin/dat Bianca en Shirley/waren betrapt/op het/stelen//
van snoep/bij de/supermarkt,/en dat/de supermarkteigenaresse/zei/dat als/één van hen//
een Snickers of een Bounty/had/gestolen,/diegene/naar/het/politiebureau/moest.//
De vriendin/zei:/"En,/wie van hen/had/een Snickers of een Bounty/gestolen/bij de/supermarkt?"//
Samantha antwoordde:/"Shirley/had/een Snickers of een Bounty/gestolen/bij/de/supermarkt."//
Shirley/zei later/dat ze/'m/had gestolen/voor/de/spanning/en om stoer/te doen.

19. (C1)
Gerben vertelde/aan/een vriend/dat/de agenten/Jan en Berend/gisteren/een huis/waren//
gaan doorzoeken,/en dat/toen/ze/terugkwamen,/de hoofdagent/zei dat/degene die//
minstens twee wapens/had/gevonden/in het huis,/promotie/zou krijgen.//
De vriend/zei:/"En,/wat/had/Berend/gevonden/in het/huis?"//
Gerben antwoordde:/"Berend/had/een steekwapen of een pistool/gevonden/in/het/huis."//
Berend/zei later/dat hij/ze/had gevonden/op/de/zolder/van het/vervallen huis.

19. Gerben answered: /" Berend / had / a knife of a gun / found / in / the / house."

19. (C2)
Els vertelde/aan/een vriendin/dat/de agentes/Inge en Deborah/gisteren/een huis/waren//
gaan doorzoeken,/en dat/toen/ze/terugkwamen,/de hoofdagente/zei dat/degene die//
minstens twee wapens/had/gevonden/in het huis,/promotie/zou krijgen.//
De vriendin/zei:/"En,/wat/had/Deborah/gevonden/in het/huis?"//
Els antwoordde:/"Deborah/had/een steekwapen of een pistool/gevonden/in/het/huis."//
Deborah/zei later/dat ze/'m/had gevonden/op/de/zolder/van het/vervallen huis.

19. (C3)
Gerben vertelde/aan/een vriend/dat/de agenten/Jan en Berend/gisteren/een huis/waren//
gaan doorzoeken,/en dat/toen/ze/terugkwamen,/de hoofdagent/zei dat/degene die//
een steekwapen of een pistool/had/gevonden/in het huis,/promotie/zou krijgen.//
De vriend/zei:/"En,/wie van hen/had/een steekwapen of een pistool/gevonden/in het/huis?"//
Gerben antwoordde:/"Berend/had/een steekwapen of een pistool/gevonden/in/het/huis."//
Berend/zei later/dat hij/ze/had gevonden/op/de/zolder/van het/vervallen huis.

19. (C4)
Els vertelde/aan/een vriendin/dat/de agentes/Inge en Deborah/gisteren/een huis/waren//
gaan doorzoeken,/en dat/toen/ze/terugkwamen,/de hoofdagente/zei dat/degene die//
een steekwapen of een pistool/had/gevonden/in het huis,/promotie/zou krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een steekwapen of een pistool/gevonden/in het/huis?"//
Els antwoordde:/"Deborah/had/een steekwapen of een pistool/gevonden/in/het/huis."//
Deborah/zei later/dat ze/'m/had gevonden/op/de/zolder/van het/vervallen huis.

20. (C1)
Sjoerd/vertelde/aan een vriend/dat Rutger en Koen/gisteren/thuis/waren gaan/lunchen//
en dat/toen ze klaar/waren,/hun vader/zei/dat degene/die/minstens twee stuks fruit/had//
genomen/bij/de lunch,/geen vitaminepil/hoefde/te nemen.//
De vriend/zei:/"En,/wat/had/Rutger/genomen/bij de/lunch?"//
Sjoerd antwoordde:/"Rutger/had/een grapefruit of een kiwi/genomen/bij/de/lunch."//
Rutger/zei later/dat hij/ze/had genomen/om/het/gevoel/te hebben/gezond te eten.

20. Sjoerd answered: / "Rutger / had / a grapefruit or a kiwi / taken / for / lunch."

20. (C2)
Maartje/vertelde/aan een vriendin/dat Stefanie en Bregje/gisteren/thuis/waren gaan/lunchen//
en dat/toen ze klaar/waren,/hun moeder/zei/dat degene/die/minstens twee stuks fruit/had//
genomen/bij/de lunch,/geen vitaminepil/hoefde/te nemen.//
De vriendin/zei:/"En,/wat/had/Stefanie/genomen/bij de/lunch?"//
Maartje antwoordde:/"Stefanie/had/een grapefruit of een kiwi/genomen/bij/de/lunch."//
Stefanie/zei later/dat ze/'m/had genomen/om/het/gevoel/te hebben/gezond te eten.

20. (C3)
Sjoerd/vertelde/aan een vriend/dat Rutger en Koen/gisteren/thuis/waren gaan/lunchen//
en dat/toen ze klaar/waren,/hun vader/zei/dat degene/die/een grapefruit of een kiwi/had//
genomen/bij/de lunch,/geen vitaminepil/hoefde/te nemen.//
De vriend/zei:/"En,/wie van hen/had/een grapefruit of een kiwi/genomen/bij de/lunch?"//
Sjoerd antwoordde:/"Rutger/had/een grapefruit of een kiwi/genomen/bij/de/lunch."//
Rutger/zei later/dat hij/ze/had genomen/om/het/gevoel/te hebben/gezond te eten.

20. (C4)
Maartje/vertelde/aan een vriendin/dat Stefanie en Bregje/gisteren/thuis/waren gaan/lunchen//
en dat/toen ze klaar/waren,/hun moeder/zei/dat degene/die/een grapefruit of een kiwi/had//
genomen/bij/de lunch,/geen vitaminepil/hoefde/te nemen.//
De vriendin/zei:/"En,/wie van hen/had/een grapefruit of een kiwi/genomen/bij de/lunch?"//
Maartje antwoordde:/"Stefanie/had/een grapefruit of een kiwi/genomen/bij/de/lunch."//
Stefanie/zei later/dat ze/'m/had genomen/om/het/gevoel/te hebben/gezond te eten.

21. (C1)
Brian vertelde/aan een vriend/dat Daniël en Ivan/gisteren/hoofddeksels/mee naar school//
moesten/nemen/voor een toneelstuk,/en dat/toen ze/op school/kwamen,/de leraar zei/dat degene//
die/meerdere hoofddeksels/had meegenomen/voor/het toneelstuk,/een/grote rol/zou krijgen.//
De vriend/zei:/"En,/wat/had/Ivan/meegenomen/voor het/toneelstuk?"//
Brian antwoordde:/"Ivan/had/een cowboyhoed of een baret/meegenomen/voor/het/toneelstuk."//
Ivan/zei later/dat hij/ze/had meegenomen/uit/de/gangkast/zonder te/vragen.

21. Brian answered: / " Ivan / had / a cowboy hat or a beret / brought / for / the / play."

21. (C2)
Kim vertelde/aan een vriendin/dat Lisanne en Merel/gisteren/hoofddeksels/mee naar school//
moesten/nemen/voor een toneelstuk,/en dat/toen ze/op school/kwamen,/de lerares zei/dat degene//
die/meerdere hoofddeksels/had meegenomen/voor/het toneelstuk,/een/grote rol/zou krijgen.//
De vriendin/zei:/"En,/wat/had/Merel/meegenomen/voor het/toneelstuk?"//
Kim antwoordde:/"Merel/had/een cowboyhoed of een baret/meegenomen/voor/het/toneelstuk."//
Merel/zei later/dat ze/'m/had meegenomen/uit/de/gangkast/zonder te/vragen.

21. (C3)
Brian vertelde/aan een vriend/dat Daniël en Ivan/gisteren/hoofddeksels/mee naar school//
moesten/nemen/voor een toneelstuk,/en dat/toen ze/op school/kwamen,/de leraar zei/dat degene//
die/een cowboyhoed of een baret/had meegenomen/voor/het toneelstuk,/een/grote rol/zou krijgen.//
De vriend/zei:/"En,/wie van hen/had/een cowboyhoed of een baret/meegenomen/voor het/toneelstuk?"//
Brian antwoordde:/"Ivan/had/een cowboyhoed of een baret/meegenomen/voor/het/toneelstuk."//
Ivan/zei later/dat hij/ze/had meegenomen/uit/de/gangkast/zonder te/vragen.

21. (C4)
Kim vertelde/aan een vriendin/dat Lisanne en Merel/gisteren/hoofddeksels/mee naar school//
moesten/nemen/voor een toneelstuk,/en dat/toen ze/op school/kwamen,/de lerares zei/dat degene//
die/een cowboyhoed of een baret/had meegenomen/voor/het toneelstuk,/een/grote rol/zou krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een cowboyhoed of een baret/meegenomen/voor het/toneelstuk?//
Kim antwoordde:/"Merel/had/een cowboyhoed of een baret/meegenomen/voor/het/toneelstuk."//
Merel/zei later/dat ze/'m/had meegenomen/uit/de/gangkast/zonder te/vragen.

22. (C1)
Lars vertelde/aan een vriend/dat Menno en Justin/tijdens/de kerstvakantie/boeken/moesten//
lezen/voor Nederlands,/en dat/toen ze/weer/op school kwamen,/de leraar zei/dat degene die//
meer dan één boek/had gelezen/in/de vakantie,/een/extra hoge/beoordeling/zou krijgen.//
De vriend/zei:/"En,/wat/had/Menno/gelezen/in de/vakantie?"//
Lars antwoordde:/"Menno/had/een dichtbundel of een roman/gelezen/in/de/vakantie."//
Menno/zei later/dat hij/ze/had gelezen/op/de/dagen/dat het toch/slecht weer was.

22. Lars answered: /"Menno / had / a book of poetry or a novel / read / during / the / vacation."

22. (C2)
Rachel vertelde/aan een vriendin/dat Sonja en Christina/tijdens/de kerstvakantie/boeken/moesten//
lezen/voor Nederlands,/en dat/toen ze/weer/op school kwamen,/de lerares zei/dat degene die//
meer dan één boek/had gelezen/in/de vakantie,/een/extra hoge/beoordeling/zou krijgen.//
De vriendin/zei:/"En,/wat/had/Sonja/gelezen/in de/vakantie?"//
Rachel antwoordde:/"Sonja/had/een dichtbundel of een roman/gelezen/in/de/vakantie."//
Sonja/zei later/dat ze/'m/had gelezen/op/de/dagen/dat het toch/slecht weer was.

22. (C3)
Lars vertelde/aan een vriend/dat Menno en Justin/tijdens/de kerstvakantie/boeken/moesten//
lezen/voor Nederlands,/en dat/toen ze/weer/op school kwamen,/de leraar zei/dat degene die//
een dichtbundel of een roman/had gelezen/in/de vakantie,/een/extra hoge/beoordeling/zou krijgen.//
De vriend/zei:/"En,/wie van hen/had/een dichtbundel of een roman/gelezen/in de/vakantie?"//
Lars antwoordde:/"Menno/had/een dichtbundel of een roman/gelezen/in/de/vakantie."//
Menno/zei later/dat hij/ze/had gelezen/op/de/dagen/dat het toch/slecht weer was.

22. (C4)
Rachel vertelde/aan een vriendin/dat Sonja en Christina/tijdens/de kerstvakantie/boeken/moesten//
lezen/voor Nederlands,/en dat/toen ze/weer/op school kwamen,/de lerares zei/dat degene die//
een dichtbundel of een roman/had gelezen/in/de vakantie,/een/extra hoge/beoordeling/zou krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een dichtbundel of een roman/gelezen/in de/vakantie?"//
Rachel antwoordde:/"Sonja/had/een dichtbundel of een roman/gelezen/in/de/vakantie."//
Sonja/zei later/dat ze/'m/had gelezen/op/de/dagen/dat het toch/slecht weer was.

23. (C1)
Jorrit vertelde/aan een vriend/dat Sjors en Roeland/vorige week/voor een fotocursus/dieren/moesten//
fotograferen,/en dat/de keer/erna/de docent zei/dat als één van hen/meer dan één dier//
had/gefotografeerd/voor de/cursus,/diegene/de/volgende opdracht/niet hoefde/te doen.//
De vriend/zei:/"En,/wat/had/Roeland/gefotografeerd/voor de/cursus?"//
Jorrit antwoordde:/"Roeland/had/een neushoorn of een giraffe/gefotografeerd/voor/de/cursus."//
Roeland/zei later/dat hij/ze/had gefotografeerd/in/het/verblijf/in de/dierentuin.

23. Jorrit answered: / "Roeland / had / a hippo or a giraffe / photographed / for / the / course.

23. (C2)
Suzanne vertelde/aan een vriendin/dat Kate en Nienke/vorige week/voor een fotocursus/dieren/moesten//
fotograferen,/en dat/de keer/erna/de docente zei/dat als één van hen/meer dan één dier//
had/gefotografeerd/voor de/cursus,/diegene/de/volgende opdracht/niet hoefde/te doen.//
De vriendin/zei:/"En,/wat/had/Nienke/gefotografeerd/voor de/cursus?"//
Suzanne antwoordde:/"Nienke/had/een neushoorn of een giraffe/gefotografeerd/voor/de/cursus."//
Nienke/zei later/dat ze/'m/had gefotografeerd/in/het/verblijf/in de/dierentuin.

23. (C3)
Jorrit vertelde/aan een vriend/dat Sjors en Roeland/vorige week/voor een fotocursus/dieren/moesten//
fotograferen,/en dat/de keer/erna/de docent zei/dat als één van hen/een neushoorn of een giraffe//
had/gefotografeerd/voor de/cursus,/diegene/de/volgende opdracht/niet hoefde/te doen.//
De vriend/zei:/"En,/wie van hen/had/een neushoorn of een giraffe/gefotografeerd/voor de/cursus?"//
Jorrit antwoordde:/"Roeland/had/een neushoorn of een giraffe/gefotografeerd/voor/de/cursus."//
Roeland/zei later/dat hij/ze/had gefotografeerd/in/het/verblijf/in de/dierentuin.

23. (C4)
Suzanne vertelde/aan een vriendin/dat Kate en Nienke/vorige week/voor een fotocursus/dieren/moesten//
fotograferen,/en dat/de keer/erna/de docente zei/dat als één van hen/een neushoorn of een giraffe//
had/gefotografeerd/voor de/cursus,/diegene/de/volgende opdracht/niet hoefde/te doen.//
De vriendin/zei:/"En,/wie van hen/had/een neushoorn of een giraffe/gefotografeerd/voor de/cursus?"//
Suzanne antwoordde:/"Nienke/had/een neushoorn of een giraffe/gefotografeerd/voor/de/cursus."//
Nienke/zei later/dat ze/'m/had gefotografeerd/in/het/verblijf/in de/dierentuin.

24. (C1)
Oscar/vertelde/aan een/vriend/dat Wesley en Roberto/gisteren/waren opgepakt/voor vernieling//
in de buurt,/en dat/de agent/zei/dat als één van hen/meerdere dingen/had/vernield//
in/de buurt,/diegene/een gevangenisstraf/boven/het hoofd/hing.//
De vriend/zei:/"En,/wat/had/Wesley/vernield/in de/buurt?"//
Oscar antwoordde:/"Wesley/had/een lantaarnpaal of een auto/vernield/in/de/buurt."//
Wesley/zei later/dat hij/ze/had vernield/om/het/respect/van de andere hooligans/te winnen.

24. Oscar answered: / "Wesley / had / a lamppost or a car / vandalized / in / the / neighborhood."

24. (C2)
Kimberly/vertelde/aan een/vriendin/dat Marja en Ramona/gisteren/waren opgepakt/voor vernieling//
in de buurt,/en dat/de agente/zei/dat als één van hen/meerdere dingen/had/vernield//
in/de buurt,/diegene/een gevangenisstraf/boven/het hoofd/hing.//
De vriendin/zei:/"En,/wat/had/Marja/vernield/in de/buurt?"//
Kimberly antwoordde:/"Marja/had/een lantaarnpaal of een auto/vernield/in/de/buurt."//
Marja/zei later/dat ze/'m/had vernield/om/het/respect/van de andere hooligans/te winnen.

24. (C3)
Oscar/vertelde/aan een/vriend/dat Wesley en Roberto/gisteren/waren opgepakt/voor vernieling//
in de buurt,/en dat/de agent/zei/dat als één van hen/een lantaarnpaal of een auto/had/vernield//
in/de buurt,/diegene/een gevangenisstraf/boven/het hoofd/hing.//
De vriend/zei:/"En,/wie van hen/had/een lantaarnpaal of een auto/vernield/in de/buurt?"//
Oscar antwoordde:/"Wesley/had/een lantaarnpaal of een auto/vernield/in/de/buurt."//
Wesley/zei later/dat hij/ze/had vernield/om/het/respect/van de andere hooligans/te winnen.

24. (C4)
Kimberly/vertelde/aan een/vriendin/dat Marja en Ramona/gisteren/waren opgepakt/voor vernieling//
in de buurt,/en dat/de agente/zei/dat als één van hen/een lantaarnpaal of een auto/had/vernield//
in/de buurt,/diegene/een gevangenisstraf/boven/het hoofd/hing.//
De vriendin/zei:/"En,/wie van hen/had/een lantaarnpaal of een auto/vernield/in de/buurt?"//
Kimberly antwoordde:/"Marja/had/een lantaarnpaal of een auto/vernield/in/de/buurt."//
Marja/zei later/dat ze/'m/had vernield/om/het/respect/van de andere hooligans/te winnen.

25. (C1)
Arnoud vertelde/aan/een vriend/dat Gijs en Carlos/gisteren/insecten/moesten eten/bij opnames//
van het TV-programma/"Fear Factor",/en dat/toen ze/terugkwamen,/hun vader/zei dat/als één van hen//
meer dan één insect/had opgegeten/in/het programma,/diegene/een held/was.//
De vriend/zei:/"En,/wat/had/Carlos/opgegeten/in het/programma?"//
Arnoud antwoordde:/"Carlos/had/een krekel of een sprinkhaan/opgegeten/in/het/programma."//
Carlos/zei later/dat hij/ze/had opgegeten/met/de/ogen/dicht en zonder/te kauwen.

25. Arnoud answered: /"Carlos / had / a cricket or a grasshopper / eaten / on / the / show."

25. (C2)
Maaike vertelde/aan/een vriendin/dat Katja en Heleen/gisteren/insecten/moesten eten/bij opnames//
van het TV-programma/"Fear Factor",/en dat/toen ze/terugkwamen,/hun moeder/zei dat/als één van hen//
meer dan één insect/had opgegeten/in/het programma,/diegene/een heldin/was.//
De vriendin/zei:/"En,/wat/had/Katja/opgegeten/in het/programma?"//
Maaike antwoordde:/"Katja/had/een krekel of een sprinkhaan/opgegeten/in/het/programma."//
Katja/zei later/dat ze/'m/had opgegeten/met/de/ogen/dicht en zonder/te kauwen.

25. (C3)
Arnoud vertelde/aan/een vriend/dat Gijs en Carlos/gisteren/insecten/moesten eten/bij opnames//
van het TV-programma/"Fear Factor",/en dat/toen ze/terugkwamen,/hun vader/zei dat/als één van hen//
een krekel of een sprinkhaan/had opgegeten/in/het programma,/diegene/een held/was.//
De vriend/zei:/"En,/wie van hen/had/een krekel of een sprinkhaan/opgegeten/in het/programma?"//
Arnoud antwoordde:/"Carlos/had/een krekel of een sprinkhaan/opgegeten/in/het/programma."//
Carlos/zei later/dat hij/ze/had opgegeten/met/de/ogen/dicht en zonder/te kauwen.

25. (C4)
Maaike vertelde/aan/een vriendin/dat Katja en Heleen/gisteren/insecten/moesten eten/bij opnames//
van het TV-programma/"Fear Factor",/en dat/toen ze/terugkwamen,/hun moeder/zei dat/als één van hen//
een krekel of een sprinkhaan/had opgegeten/in/het programma,/diegene/een heldin/was.//
De vriendin/zei:/"En,/wie van hen/had/een krekel of een sprinkhaan/opgegeten/in het/programma?"//
Maaike antwoordde:/"Katja/had/een krekel of een sprinkhaan/opgegeten/in/het/programma."//
Katja/zei later/dat ze/'m/had opgegeten/met/de/ogen/dicht en zonder/te kauwen.

26. (C1)
Erik/vertelde/aan/een vriend/dat Jesse en Lucas/gisteren hadden/meegedaan/aan een//
circus-workshop,/en dat/aan het/einde/de circusdirecteur/had gezegd/dat degene/die//
minstens twee circusdieren/had bereden/bij/de/workshop,/een certificaat/kreeg.//
De vriend/zei:/"En,/wat/had/Lucas/bereden/bij de/workshop?//
Erik antwoordde:/"Lucas/had/een olifant of een kameel/bereden/bij/de/workshop."//
Lucas/zei later/dat hij/ze/had bereden/om/het/gevoel/van een echte artiest/te krijgen.

26. Erik answered: /"Lucas / had / an elephant or a camel / ridden / during / the / workshop."

26. (C2)
Hilde/vertelde/aan/een vriendin/dat Maud en Noortje/gisteren hadden/meegedaan/aan een//
circus-workshop,/en dat/aan het/einde/de circusdirectrice/had gezegd/dat degene/die//
minstens twee circusdieren/had bereden/bij/de/workshop,/een certificaat/kreeg.//
De vriendin/zei:/"En,/wat/had/Noortje/bereden/bij de/workshop?//
Hilde antwoordde:/"Noortje/had/een olifant of een kameel/bereden/bij/de/workshop."//
Noortje/zei later/dat ze/'m/had bereden/om/het/gevoel/van een echte artiest/te krijgen.

26. (C3)
Erik/vertelde/aan/een vriend/dat Jesse en Lucas/gisteren hadden/meegedaan/aan een//
circus-workshop,/en dat/aan het/einde/de circusdirecteur/had gezegd/dat degene/die//
een olifant of een kameel/had bereden/bij/de/workshop,/een certificaat/kreeg.//
De vriend/zei:/"En,/wie van hen/had/een olifant of een kameel/bereden/bij de/workshop?//
Erik antwoordde:/"Lucas/had/een olifant of een kameel/bereden/bij/de/workshop."//
Lucas/zei later/dat hij/ze/had bereden/om/het/gevoel/van een echte artiest/te krijgen.

26. (C4)
Hilde/vertelde/aan/een vriendin/dat Maud en Noortje/gisteren hadden/meegedaan/aan een//
circus-workshop,/en dat/aan het/einde/de circusdirectrice/had gezegd/dat degene/die//
een olifant of een kameel/had bereden/bij/de/workshop,/een certificaat/kreeg.//
De vriendin/zei:/"En,/wie van hen/had/een olifant of een kameel/bereden/bij de/workshop?//
Hilde antwoordde:/"Noortje/had/een olifant of een kameel/bereden/bij/de/workshop."//
Noortje/zei later/dat ze/'m/had bereden/om/het/gevoel/van een echte artiest/te krijgen.

27. (C1)
David/vertelde/aan een vriend dat/Ferry en Mario/vorige week/naar/de VS/waren geweest,/en//
dat tijdens/de/terugvlucht/de steward/zei/dat als/één van hen/meer dan één elektronisch apparaat//
had meegenomen/uit/de VS,/diegene/invoerrechten/moest betalen.//
De vriend/zei:/"En,/wat/had/Ferry/meegenomen/uit de/VS?"//
David antwoordde:/"Ferry/had/een telefoon of een i-pod/meegenomen/uit/de/VS."//
Ferry/zei later/dat hij/ze/had meegenomen/voor/een/kennis/die ervoor/had betaald.

27. David answered: /"Ferry / had / a phone or an i-pod/ brought / from / the / US.

27. (C2)
Hanna/vertelde/aan een vriendin dat/Birgit en Loes/vorige week/naar/de VS/waren geweest,/en//
dat tijdens/de/terugvlucht/de stewardess/zei/dat als/één van hen/meer dan één elektronisch apparaat//
had meegenomen/uit/de VS,/diegene/invoerrechten/moest betalen.//
De vriendin/zei:/"En,/wat/had/Birgit/meegenomen/uit de/VS?"//
Hanna antwoordde:/"Birgit/had/een telefoon of een i-pod/meegenomen/uit/de/VS."//
Birgit/zei later/dat ze/'m/had meegenomen/voor/een/kennis/die ervoor/had betaald.

27. (C3)
David/vertelde/aan een vriend dat/Ferry en Mario/vorige week/naar/de VS/waren geweest,/en//
dat tijdens/de/terugvlucht/de steward/zei/dat als/één van hen/een telefoon of een i-pod//
had meegenomen/uit/de VS,/diegene/invoerrechten/moest betalen.//
De vriend/zei:/"En,/wie van hen/had/een telefoon of een i-pod/meegenomen/uit de/VS?"//
David antwoordde:/"Ferry/had/een telefoon of een i-pod/meegenomen/uit/de/VS."//
Ferry/zei later/dat hij/ze/had meegenomen/voor/een/kennis/die ervoor/had betaald.

27. (C4)
Hanna/vertelde/aan een vriendin dat/Birgit en Loes/vorige week/naar/de VS/waren geweest,/en//
dat tijdens/de/terugvlucht/de stewardess/zei/dat als/één van hen/een telefoon of een i-pod//
had meegenomen/uit/de VS,/diegene/invoerrechten/moest betalen.//
De vriendin/zei:/"En,/wie van hen/had/een telefoon of een i-pod/meegenomen/uit de/VS?"//
Hanna antwoordde:/"Birgit/had/een telefoon of een i-pod/meegenomen/uit/de/VS."//
Birgit/zei later/dat ze/'m/had meegenomen/voor/een/kennis/die ervoor/had betaald.

28. (C1)
Pim vertelde/aan/een/vriend/dat Chris en Arno/waren gaan/fitnessen,/en dat/toen ze//
terugkwamen,/hun huisgenoot/zei/dat als één van hen/meer dan één versnapering/had gegeten//
na/de training,/diegene/net zo goed/niet/had kunnen/gaan/sporten.//
De vriend/zei:/"En,/wat/had/Arno/gegeten/na de/training?"//
Pim antwoordde:/"Arno/had/een koek of een chocoladereep/gegeten/na/de/training."//
Arno/zei later/dat hij/ze/had gegeten/om/de/suikers/in het bloed/aan te vullen.

28. Pim answered: / "Arno / had / a cookie or a chocolate bar/ eaten / after / practice."

28. (C2)
Karien vertelde/aan/een/vriendin/dat Marieke en Kirsten/waren gaan/fitnessen,/en dat/toen ze//
terugkwamen,/hun huisgenootje/zei/dat als één van hen/meer dan één versnapering/had gegeten//
na/de training,/diegene/net zo goed/niet/had kunnen/gaan/sporten.//
De vriendin/zei:/"En,/wat/had/Kirsten/gegeten/na de/training?"//
Karien antwoordde:/"Kirsten/had/een koek of een chocoladereep/gegeten/na/de/training."//
Kirsten/zei later/dat ze/'m/had gegeten/om/de/suikers/in het bloed/aan te vullen.

28. (C3)
Pim vertelde/aan/een/vriend/dat Chris en Arno/waren gaan/fitnessen,/en dat/toen ze//
terugkwamen,/hun huisgenoot/zei/dat als één van hen/een koek of een chocoladereep/had gegeten//
na/de training,/diegene/net zo goed/niet/had kunnen/gaan/sporten.//
De vriend/zei:/"En,/wie van hen/had/een koek of een chocoladereep/gegeten/na de/training?"//
Pim antwoordde:/"Arno/had/een koek of een chocoladereep/gegeten/na/de/training."//
Arno/zei later/dat hij/ze/had gegeten/om/de/suikers/in het bloed/aan te vullen.

28. (C4)
Karien vertelde/aan/een/vriendin/dat Marieke en Kirsten/waren gaan/fitnessen,/en dat/toen ze//
terugkwamen,/hun huisgenootje/zei/dat als één van hen/een koek of een chocoladereep/had gegeten//
na/de training,/diegene/net zo goed/niet/had kunnen/gaan/sporten.//
De vriendin/zei:/"En,/wie van hen/had/een koek of een chocoladereep/gegeten/na de/training?"//
Karien antwoordde:/"Kirsten/had/een koek of een chocoladereep/gegeten/na/de/training."//
Kirsten/zei later/dat ze/'m/had gegeten/om/de/suikers/in het bloed/aan te vullen.

29. (C1)
Victor/vertelde aan/een vriend dat/Boris en Jacco/gisteren/langs de deuren/waren gegaan/om//
huishoudelijke artikelen/te verkopen,/en dat/toen/ze/terugkwamen,/hun baas zei/dat degene die//
minstens twee artikelen/had verkocht/op/die dag,/een bonus/zou/krijgen.//
De vriend/zei:/"En,/wat/had/Boris/verkocht/op die/dag?"//
Victor antwoordde:/"Boris/had/een strijkbout of een stofzuiger/verkocht/op/die/dag."//
Boris/zei later/dat hij/ze/had verkocht/aan/een/vrouwtje/van minstens/zeventig.

29. Vicotr answered: / "Boris / had / an iron or a vacuum cleaner / sold / on / that / day."

29. (C2)
Daphne/vertelde aan/een vriendin dat/Tessa en Nina/gisteren/langs de deuren/waren gegaan/om//
huishoudelijke artikelen/te verkopen,/en dat/toen/ze/terugkwamen,/hun bazin zei/dat degene die//
minstens twee artikelen/had verkocht/op/die dag,/een bonus/zou/krijgen.//
De vriendin/zei:/"En,/wat/had/Tessa/verkocht/op die/dag?"//
Daphne antwoordde:/"Tessa/had/een strijkbout of een stofzuiger/verkocht/op/die/dag."//
Tessa/zei later/dat ze/'m/had verkocht/aan/een/vrouwtje/van minstens/zeventig.

29. (C3)
Victor/vertelde aan/een vriend dat/Boris en Jacco/gisteren/langs de deuren/waren gegaan/om//
huishoudelijke artikelen/te verkopen,/en dat/toen/ze/terugkwamen,/hun baas zei/dat degene die//
een strijkbout of een stofzuiger/had verkocht/op/die dag,/een bonus/zou/krijgen.//
De vriend/zei:/"En,/wie van hen/had/een strijkbout of een stofzuiger/verkocht/op die/dag?"//
Victor antwoordde:/"Boris/had/een strijkbout of een stofzuiger/verkocht/op/die/dag."//
Boris/zei later/dat hij/ze/had verkocht/aan/een/vrouwtje/van minstens/zeventig.

29. (C4)
Daphne/vertelde aan/een vriendin dat/Tessa en Nina/gisteren/langs de deuren/waren gegaan/om//
huishoudelijke artikelen/te verkopen,/en dat/toen/ze/terugkwamen,/hun bazin zei/dat degene die//
een strijkbout of een stofzuiger/had verkocht/op/die dag,/een bonus/zou/krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een strijkbout of een stofzuiger/verkocht/op die/dag?"//
Daphne antwoordde:/"Tessa/had/een strijkbout of een stofzuiger/verkocht/op/die/dag."//
Tessa/zei later/dat ze/'m/had verkocht/aan/een/vrouwtje/van minstens/zeventig.

30. (C1)
Aart/vertelde/aan een vriend/dat/Olaf en Sven/gisteren/naar de Gamma/waren/geweest//
en dat/er een actie/was/dat als je/meerdere apparaten/had gekocht,/je een/bon//
zou krijgen/waarmee je/de helft/van het geld/terug/kon krijgen.//
De vriend/zei:/"En,/wat/had/Olaf/gekocht/bij de/Gamma?"//
Aart antwoordde:/"Olaf/had/een schuurmachine of een boor/gekocht/bij/de/Gamma."//
Olaf/zei later/dat hij/ze/had gekocht/met/het/idee/het huis/eens wat op te knappen.

30. Aart answered: / "Olaf / had / a sanding machine or a drill / bought / at / the / Gamma."

30. (C2)
Dorine/vertelde/aan een vriendin/dat/Ella en Lotte/gisteren/naar de Gamma/waren/geweest//
en dat/er een actie/was/dat als je/meerdere apparaten/had gekocht,/je een/bon//
zou krijgen/waarmee je/de helft/van het geld/terug/kon krijgen.//
De vriendin/zei:/"En,/wat/had/Lotte/gekocht/bij de/Gamma?"//
Dorine antwoordde:/"Lotte/had/een schuurmachine of een boor/gekocht/bij/de/Gamma."//
Lotte/zei later/dat ze/'m/had gekocht/met/het/idee/het huis/eens wat op te knappen.

30. (C3)
Aart/vertelde/aan een vriend/dat/Olaf en Sven/gisteren/naar de Gamma/waren/geweest//
en dat/er een actie/was/dat als je/een schuurmachine of een boor/had gekocht,/je een/bon//
zou krijgen/waarmee je/de helft/van het geld/terug/kon krijgen.//
De vriend/zei:/"En,/wie van hen/had/een schuurmachine of een boor/gekocht/bij de/Gamma?"//
Aart antwoordde:/"Olaf/had/een schuurmachine of een boor/gekocht/bij/de/Gamma."//
Olaf/zei later/dat hij/ze/had gekocht/met/het/idee/het huis/eens wat op te knappen.

30. (C4)
Dorine/vertelde/aan een vriendin/dat/Ella en Lotte/gisteren/naar de Gamma/waren/geweest//
en dat/er een actie/was/dat als je/een schuurmachine of een boor/had gekocht,/je een/bon//
zou krijgen/waarmee je/de helft/van het geld/terug/kon krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een schuurmachine of een boor/gekocht/bij de/Gamma?"//
Dorine antwoordde:/"Lotte/had/een schuurmachine of een boor/gekocht/bij/de/Gamma."//
Lotte/zei later/dat ze/'m/had gekocht/met/het/idee/het huis/eens wat op te knappen.

31. (C1)
Winston/vertelde/aan een vriend/dat/Leo en Rudolf/gisteren/op/een motorbeurs/waren/geweest,//
en dat/er/een actie was/dat als/je/meer dan één motor/had proefgereden/op/de beurs,/je//
een gratis/sleutelhanger/kreeg.//
De vriend/zei:/"En,/wat/had/Rudolf/proefgereden/op de/beurs?"//
Winston antwoordde:/"Rudolf/had/een Harley of een BMW/proefgereden/op/de/beurs."//
Rudolf/zei later/dat hij/ze/had proefgereden/om/het/geluid/te kunnen/horen.

31. Winston answered: / "Rudolf / had / a Harley or a BMW / tried out / on / the / fair."

31. (C2)
Nikki/vertelde/aan een vriendin/dat/Jet en Wilma/gisteren/op/een motorbeurs/waren/geweest,//
en dat/er/een actie was/dat als/je/meer dan één motor/had proefgereden/op/de beurs,/je//
een gratis/sleutelhanger/kreeg.//
De vriendin/zei:/"En,/wat/had/Wilma/proefgereden/op de/beurs?"//
Nikki antwoordde:/"Wilma/had/een Harley of een BMW/proefgereden/op/de/beurs."//
Wilma/zei later/dat ze/'m/had proefgereden/om/het/geluid/te kunnen/horen.

31. (C3)
Winston/vertelde/aan een vriend/dat/Leo en Rudolf/gisteren/op/een motorbeurs/waren/geweest,//
en dat/er/een actie was/dat als/je/een Harley of een BMW/had proefgereden/op/de beurs,/je//
een gratis/sleutelhanger/kreeg.//
De vriend/zei:/"En,/wie van hen/had/een Harley of een BMW/proefgereden/op de/beurs?"//
Winston antwoordde:/"Rudolf/had/een Harley of een BMW/proefgereden/op/de/beurs."//
Rudolf/zei later/dat hij/ze/had proefgereden/om/het/geluid/te kunnen/horen.

31. (C4)
Nikki/vertelde/aan een vriendin/dat/Jet en Wilma/gisteren/op/een motorbeurs/waren/geweest,//
en dat/er/een actie was/dat als/je/een Harley of een BMW/had proefgereden/op/de beurs,/je//
een gratis/sleutelhanger/kreeg.//
De vriendin/zei:/"En,/wie van hen/had/een Harley of een BMW/proefgereden/op de/beurs?"//
Nikki antwoordde:/"Wilma/had/een Harley of een BMW/proefgereden/op/de/beurs."//
Wilma/zei later/dat ze/'m/had proefgereden/om/het/geluid/te kunnen/horen.

32. (C1)
Pascal/vertelde/aan/een vriend/dat/Ronald en Jelmer/gisteren aan/een/kookwedstrijd//
hadden/meegedaan,/en dat/de wedstrijdleider/aan het einde/zei/dat degene die//
meerdere gerechten/had gemaakt/bij de/wedstrijd,/gediskwalificeerd/zou/worden.//
De vriend/zei:/"En,/wat/had/Jelmer/gemaakt/bij de/wedstrijd?"//
Pascal antwoordde:/"Jelmer/had/een soep of een ovenschotel/gemaakt/bij/de/wedstrijd."//
Jelmer/zei later/dat hij/ze/had gemaakt/met/een/aantal/geheime/ingrediënten.

32. Pascal answered: / "Jelmer / had / a soup or an oven dish / made / at / the / competition.

32. (C2)
Petra/vertelde/aan/een vriendin/dat/Mariska en Jolien/gisteren aan/een/kookwedstrijd//
hadden/meegedaan,/en dat/de wedstrijdleidster/aan het einde/zei/dat degene die//
meerdere gerechten/had gemaakt/bij de/wedstrijd,/gediskwalificeerd/zou/worden.//
De vriendin/zei:/"En,/wat/had/Jolien/gemaakt/bij de/wedstrijd?"//
Petra antwoordde:/"Jolien/had/een soep of een ovenschotel/gemaakt/bij/de/wedstrijd."//
Jolien/zei later/dat ze/'m/had gemaakt/met/een/aantal/geheime/ingrediënten.

32. (C3)
Pascal/vertelde/aan/een vriend/dat/Ronald en Jelmer/gisteren aan/een/kookwedstrijd//
hadden/meegedaan,/en dat/de wedstrijdleider/aan het einde/zei/dat degene die//
een soep of een ovenschotel/had gemaakt/bij de/wedstrijd,/gediskwalificeerd/zou/worden.//
De vriend/zei:/"En,/wie van hen/had/een soep of een ovenschotel/gemaakt/bij de/wedstrijd?"//
Pascal antwoordde:/"Jelmer/had/een soep of een ovenschotel/gemaakt/bij/de/wedstrijd."//
Jelmer/zei later/dat hij/ze/had gemaakt/met/een/aantal/geheime/ingrediënten.

32. (C4)
Petra/vertelde/aan/een vriendin/dat/Mariska en Jolien/gisteren aan/een/kookwedstrijd//
hadden/meegedaan,/en dat/de wedstrijdleidster/aan het einde/zei/dat degene die//
een soep of een ovenschotel/had gemaakt/bij de/wedstrijd,/gediskwalificeerd/zou/worden.//
De vriendin/zei:/"En,/wie van hen/had/een soep of een ovenschotel/gemaakt/bij de/wedstrijd?"//
Petra antwoordde:/"Jolien/had/een soep of een ovenschotel/gemaakt/bij/de/wedstrijd."//
Jolien/zei later/dat ze/'m/had gemaakt/met/een/aantal/geheime/ingrediënten.

33. (C1)
Thijs vertelde/aan een vriend/dat Edwin en Rogier/vorige week/met/het vliegtuig/op//
vakantie/gingen,/en dat/toen ze/in het/vliegtuig/zaten,/de steward/omriep/dat degene/die//
meer dan één tas/had meegenomen/in/het vliegtuig,/moest bijbetalen.//
De vriend/zei:/"En,/wat/had/Rogier/meegenomen/in het/vliegtuig?"//
Thijs antwoordde:/"Rogier/had/een rugtas of een koffer/meegenomen/in/het/vliegtuig."//
Rogier/zei later/dat hij/ze/had meegenomen/om/een/oogje/in het zeil/te kunnen houden.

33. Thijs answered: / "Rogier / had / a backpack or a suitcase / brought / on / the / plane."

33. (C2)
Victoria vertelde/aan een vriendin/dat Eline en Iris/vorige week/met/het vliegtuig/op//
vakantie/gingen,/en dat/toen ze/in het/vliegtuig/zaten,/de stewardess/omriep/dat degene/die//
meer dan één tas/had meegenomen/in/het vliegtuig,/moest bijbetalen.//
De vriendin/zei:/"En,/wat/had/Iris/meegenomen/in het/vliegtuig?"//
Victoria antwoordde:/"Iris/had/een rugtas of een koffer/meegenomen/in/het/vliegtuig."//
Iris/zei later/dat ze/'m/had meegenomen/om/een/oogje/in het zeil/te kunnen houden.

33. (C3)
Thijs vertelde/aan een vriend/dat Edwin en Rogier/vorige week/met/het vliegtuig/op//
vakantie/gingen,/en dat/toen ze/in het/vliegtuig/zaten,/de steward/omriep/dat degene/die//
een rugtas of een koffer/had meegenomen/in/het vliegtuig,/moest bijbetalen.//
De vriend/zei:/"En,/wie van hen/had/een rugtas of een koffer/meegenomen/in het/vliegtuig?"//
Thijs antwoordde:/"Rogier/had/een rugtas of een koffer/meegenomen/in/het/vliegtuig."//
Rogier/zei later/dat hij/ze/had meegenomen/om/een/oogje/in het zeil/te kunnen houden.

33. (C4)
Victoria vertelde/aan een vriendin/dat Eline en Iris/vorige week/met/het vliegtuig/op//
vakantie/gingen,/en dat/toen ze/in het/vliegtuig/zaten,/de stewardess/omriep/dat degene/die//
een rugtas of een koffer/had meegenomen/in/het vliegtuig,/moest bijbetalen.//
De vriendin/zei:/"En,/wie van hen/had/een rugtas of een koffer/meegenomen/in het/vliegtuig?"//
Victoria antwoordde:/"Iris/had/een rugtas of een koffer/meegenomen/in/het/vliegtuig."//
Iris/zei later/dat ze/'m/had meegenomen/om/een/oogje/in het zeil/te kunnen houden.

34. (C1)
Wouter/vertelde aan/een vriend/dat Reinout en Thomas/vorige week/promotiemateriaal/waren gaan//
ontwerpen/voor/een/bedrijf,/en dat/de bedrijfsleider/achteraf zei/dat als één van hen//
meerdere materialen/had ontworpen/voor/het bedrijf,/diegene/extra/geld/zou krijgen.//
De vriend/zei:/"En,/wat/had/Reinout/ontworpen/voor het/bedrijf?"//
Wouter antwoordde:/"Reinout/had/een poster of een folder/ontworpen/voor/het/bedrijf."//
Reinout/zei later/dat hij/ze/had ontworpen/in/de/huisstijl/van het/bedrijf.

34. Wouter answered: / "Reinout / had / a poster or a folder/ designed / for / the / company."

34. (C2)
Tamara/vertelde aan/een vriendin/dat Marjet en Floortje/vorige week/promotiemateriaal/waren gaan//
ontwerpen/voor/een/bedrijf,/en dat/de bedrijfsleidster/achteraf zei/dat als één van hen//
meerdere materialen/had ontworpen/voor/het bedrijf,/diegene/extra/geld/zou krijgen.//
De vriendin/zei:/"En,/wat/had/Marjet/ontworpen/voor het/bedrijf?"//
Tamara antwoordde:/"Marjet/had/een poster of een folder/ontworpen/voor/het/bedrijf."//
Marjet/zei later/dat ze/'m/had ontworpen/in/de/huisstijl/van het/bedrijf.

34. (C3)
Wouter/vertelde aan/een vriend/dat Reinout en Thomas/vorige week/promotiemateriaal/waren gaan//
ontwerpen/voor/een/bedrijf,/en dat/de bedrijfsleider/achteraf zei/dat als één van hen//
een poster of een folder/had ontworpen/voor/het bedrijf,/diegene/extra/geld/zou krijgen.//
De vriend/zei:/"En,/wie van hen/had/een poster of een folder/ontworpen/voor het/bedrijf?"//
Wouter antwoordde:/"Reinout/had/een poster of een folder/ontworpen/voor/het/bedrijf."//
Reinout/zei later/dat hij/ze/had ontworpen/in/de/huisstijl/van het/bedrijf.

34. (C4)
Tamara/vertelde aan/een vriendin/dat Marjet en Floortje/vorige week/promotiemateriaal/waren gaan//
ontwerpen/voor/een/bedrijf,/en dat/de bedrijfsleidster/achteraf zei/dat als één van hen//
een poster of een folder/had ontworpen/voor/het bedrijf,/diegene/extra/geld/zou krijgen.//
De vriendin/zei:/"En,/wie van hen/had/een poster of een folder/ontworpen/voor het/bedrijf?"//
Tamara antwoordde:/"Marjet/had/een poster of een folder/ontworpen/voor/het/bedrijf."//
Marjet/zei later/dat ze/'m/had ontworpen/in/de/huisstijl/van het/bedrijf.

35. (C1)
Hans/vertelde/aan/een vriend/dat/Sander en Joram/gisteren/hadden/meegedaan/aan de//
plaatselijke/Bingo,/en dat/als je/meerdere prijzen/had/gewonnen/bij/de Bingo,//
je/een liedje/moest/zingen.//
De vriend/zei:/"En,/wat/had/Joram/gewonnen/bij de/Bingo?"//
Hans antwoordde:/"Joram/had/een appeltaart of een cake/gewonnen/bij/de/Bingo."//
Joram/zei later/dat hij/ze/had gewonnen/door/de/nummers/goed in de gaten/te houden.

35. Hans answered: / "Joram / had /an apple pie or a cake/ won / at / the / Bingo."

35. (C2)
Annemiek/vertelde/aan/een vriendin/dat/Claire en Geertje/gisteren/hadden/meegedaan/aan de//
plaatselijke/Bingo,/en dat/als je/meerdere prijzen/had/gewonnen/bij/de Bingo,//
je/een liedje/moest/zingen.//
De vriendin/zei:/"En,/wat/had/Geertje/gewonnen/bij de/Bingo?"//
Annemiek antwoordde:/"Geertje/had/een appeltaart of een cake/gewonnen/bij/de/Bingo."//
Geertje/zei later/dat ze/'m/had gewonnen/door/de/nummers/goed in de gaten/te houden.

35. (C3)
Hans/vertelde/aan/een vriend/dat/Sander en Joram/gisteren/hadden/meegedaan/aan de//
plaatselijke/Bingo,/en dat/als je/een appeltaart of een cake/had/gewonnen/bij/de Bingo,//
je/een liedje/moest/zingen.//
De vriend/zei:/"En,/wie van hen/had/een appeltaart of een cake/gewonnen/bij de/Bingo?"//
Hans antwoordde:/"Joram/had/een appeltaart of een cake/gewonnen/bij/de/Bingo."//
Joram/zei later/dat hij/ze/had gewonnen/door/de/nummers/goed in de gaten/te houden.

35. (C4)
Annemiek/vertelde/aan/een vriendin/dat/Claire en Geertje/gisteren/hadden/meegedaan/aan de//
plaatselijke/Bingo,/en dat/als je/een appeltaart of een cake/had/gewonnen/bij/de Bingo,//
je/een liedje/moest/zingen.//
De vriendin/zei:/"En,/wie van hen/had/een appeltaart of een cake/gewonnen/bij de/Bingo?"//
Annemiek antwoordde:/"Geertje/had/een appeltaart of een cake/gewonnen/bij/de/Bingo."//
Geertje/zei later/dat ze/'m/had gewonnen/door/de/nummers/goed in de gaten/te houden.

36. (C1)
Eddy vertelde/aan/een vriend/dat Jelle en Harmen/gisteren/voor tekenles/de stad in/waren//
gestuurd/om gebouwen/na te tekenen,/en dat toen/ze terugkwamen/de leraar/zei/dat degene/die//
minstens twee gebouwen/had/nagetekend/in de stad,/naar huis/mocht.//
De vriend/zei:/"En,/wat/had/Jelle/nagetekend/in de/stad?"//
Eddy antwoordde:/"Jelle/had/een kerktoren of een moskee/nagetekend/in/de/stad."//
Jelle/zei later/dat hij/ze/had nagetekend/met/een/potlood/dat op straat/lag.

36. Eddy answered: / "Jelle / had / a church tower or a mosque / drawn / in / the / city.

36. (C2)
Sylvie vertelde/aan/een vriendin/dat Tina en Jolijn/gisteren/voor tekenles/de stad in/waren//
gestuurd/om gebouwen/na te tekenen,/en dat toen/ze terugkwamen/de lerares/zei/dat degene/die//
minstens twee gebouwen/had/nagetekend/in de stad,/naar huis/mocht.//
De vriendin/zei:/"En,/wat/had/Tina/nagetekend/in de/stad?"//
Sylvie antwoordde:/"Tina/had/een kerktoren of een moskee/nagetekend/in/de/stad."//
Tina/zei later/dat ze/'m/had nagetekend/met/een/potlood/dat op straat/lag.

36. (C3)
Eddy vertelde/aan/een vriend/dat Jelle en Harmen/gisteren/voor tekenles/de stad in/waren//
gestuurd/om gebouwen/na te tekenen,/en dat toen/ze terugkwamen/de leraar/zei/dat degene/die//
een kerktoren of een moskee/had/nagetekend/in de stad,/naar huis/mocht.//
De vriend/zei:/"En,/wie van hen/had/een kerktoren of een moskee/nagetekend/in de/stad?"//
Eddy antwoordde:/"Jelle/had/een kerktoren of een moskee/nagetekend/in/de/stad."//
Jelle/zei later/dat hij/ze/had nagetekend/met/een/potlood/dat op straat/lag.

36. (C4)
Sylvie vertelde/aan/een vriendin/dat Tina en Jolijn/gisteren/voor tekenles/de stad in/waren//
gestuurd/om gebouwen/na te tekenen,/en dat toen/ze terugkwamen/de lerares/zei/dat degene/die//
een kerktoren of een moskee/had/nagetekend/in de stad,/naar huis/mocht.//
De vriendin/zei:/"En,/wie van hen/had/een kerktoren of een moskee/nagetekend/in de/stad?"//
Sylvie antwoordde:/"Tina/had/een kerktoren of een moskee/nagetekend/in/de/stad."//
Tina/zei later/dat ze/'m/had nagetekend/met/een/potlood/dat op straat/lag.

APPENDIX 8

Items used in Experiments 10 and 11 (chapter 7)

C1 = wh focus C2 = wh non-focus C3 = yes/no sentential, C4 = yes/no polar

For items 2-8 I only translate condition 1, as the other conditions are easily derived.

1. Er waren vijf aanwezigen op de borrel van de opleiding Keltisch op de universiteit. Er waren verschillende drankjes verkrijgbaar.

Sander is een student. Hij dronk bier.
Tom is ook een student. Hij dronk ook bier.
Eric is een docent. Hij dronk wijn.
Martin is ook een docent. Hij dronk appelsap.
Frans is een student. Hij dronk bier.

1. Five people were present at the get-together of the Celtic language studies program at the University. Several drinks were available.

Sander is a student. He drank beer.
Tom is also a student. He also drank beer.
Eric is a professor. He drank wine.
Martin is a professor too. He drank apple juice.
Frans is a student. He drank beer.

(C1) A: "Hoeveel van de studenten dronken bier?"
B: "De meeste studenten dronken bier."
A: "How many of the students drank beer?"
B: "Most students drank beer."

(C2) A: "Wat dronken de meeste studenten?"
B: "De meeste studenten dronken bier."
A: "What did most students drink?"
B: "Most students drank beer.

(C3) A: "Dronken de meeste studenten bier?"
B: "De meeste studenten dronken bier."
A: "Did most student drink beer?"
B: "Most students drank beer.

(C4) Dronken de meeste studenten bier? ja / nee
Did most student drink beer? yes / no

2. In het teken van de Week van het Fruit mochten de zes overblijfkinderen op een school fruit uitzoeken uit een fruitmand.
Freek nam een peer.
Johanna nam een appel.
Karlijn nam een appel.
Jeffrey nam een banaan.
Daphne nam een appel.
Marie nam een appel.

2. Because it is the Week of Fruit six schoolchildren get to pick fruit from a fruit basket.
Freek took a pear.
Johanna took an apple.
Karlijn took a apple.
Jeffrey took a banana.
Daphne took an apple.
Marie took an apple.

(C1) A: "Hoeveel van de meisjes namen een appel?"
B: "De meeste meisjes namen een appel."
A: "How many of the girls took an apple?"
B: "Most girls took an apple."

(C2) A: "Wat namen de meeste meisjes?"
B: "De meeste meisjes namen een appel."

(C3) A: "Namen de meeste meisjes een appel?"
B: "De meeste meisjes namen een appel."

(C4) Namen de meeste meisjes een appel? ja / nee

3. Drie kinderen en twee begeleiders waren op de kermis. Ze stopten bij een snoepkraampje.
Het eerste kind nam een zuurstok.
Het tweede kind nam ook een zuurstok.
Het derde kind nam ook een zuurstok.
De éne begeleider nam een wijnbal.
De andere begeleider nam een suikerspin.

3. Three children and two adults were at the fair. They stopped at a candy stand.
The first child took a stick of rock.
The second child also took a stick of rock.
The third child also took a stick of rock.
The first adult took a wine coloured sweet.
The second adult took cotton candy.

(C1) A: “Hoeveel van de kinderen namen een zuurstok?”
B: “De meeste kinderen namen een zuurstok.”
A: “How many of the children took a stick of rock?”
B: “Most children took a stick of rock.”

(C2) A: “Wat namen de meeste kinderen?”
B: “De meeste kinderen namen een zuurstok.”

(C3) A: “Namen de meeste kinderen een zuurstok?”
B: “De meeste kinderen namen een zuurstok.”

(C4) Namen de meeste kinderen een zuurstok? ja / nee

4. Zeven vrienden en vriendinnen gingen barbecuen bij één van hen thuis. Ze zouden allemaal iets meenemen.
Jan had vlees meegenomen.
Wendy had brood meegenomen.
Franc had vlees meegenomen.
Anja had een salade meegenomen.
Niels had ook vlees meegenomen.
Femke had een fles wijn meegenomen.
Joep had ook vlees meegenomen.

4. Seven friends (male and female) were having a barbecue at one of their houses.
They all agreed to bring something.
Jan had brought meat.
Wendy had brought bread.
Franc had brought meat.
Anja had brought a salad.
Niels had also brought meat.
Femke had brought a bottle of wine.
Joep had also brought meat.

(C1) A: “Hoeveel van de mannen hadden vlees meegenomen?”
B: “De meeste mannen hadden vlees meegenomen.”
A: “How many of the men had brought meat?”
B: “Most men had brought meat.”

(C2) A: “Wat hadden de meeste mannen meegenomen?”
B: “De meeste mannen hadden vlees meegenomen.”

(C3) A: “Hadden de meeste mannen vlees meegenomen?”
B: “De meeste mannen hadden vlees meegenomen.”

(C4) Hadden de meeste mannen vlees meegenomen? ja / nee

5. Drie directeuren waren een avondje op stap met hun secretaresses.
Peter droeg een pak. Angela, zijn secretaresse, droeg een jurk.
Wim droeg ook een pak. Coby, zijn secretaresse, droeg een mantelpakje.
Julia, de secretaresse van Boudewijn, droeg een jurk. Boudewijn zelf droeg een pak.

5. Three CEOs had a night out with their (female) secretaries.
Peter was wearing a suit. Angela, his secretary, was wearing a dress.
Wim was also wearing a suit. Coby, his secretary, was wearing a women's suit.
Julia, Boudewijn's secretary, was wearing a dress. Boudewijn himself was wearing a suit.

(C1) A: "Hoeveel van de directeuren droegen een pak?"
B: "De meeste directeuren droegen een pak."
A: "How many of the CEOs were wearing a suit?"
B: "Most CEOs were wearing a suit."

(C2) A: "Wat droegen de meeste directeuren?"
B: "De meeste directeuren droegen een pak."

(C3) A: "Droegen de meeste directeuren een pak?"
B: De meeste directeuren droegen een pak."

(C4) Droegen de meeste directeuren een pak? ja / nee

6. Drie Amerikanen en vier Engelsen zaten in een café. Ze waren de enige gasten.
De eerste Engelsman nam een koffie.
De tweede Engelsman nam ook een koffie.
De derde Engelsman nam een biertje.
De vierde Engelsman nam een koffie.
De eerste Amerikaan nam een biertje.
De tweede Amerikaan nam ook een biertje.
De derde Amerikaan nam ook een biertje.

6. Three Americans and four Englishmen were sitting in a café. They were the only guests.
The first Englishman took a coffee.
The second Englishman also took a coffee.
The third Englishman took a beer.
The fourth Englishman took a coffee.
The first American took a beer.
The second American also took a beer.
The third American also took a beer.

(C1) A: "Hoeveel van de Amerikanen namen een biertje?"
B: "De meeste Amerikanen namen een biertje."
A: "How many of the Americans took a beer?"
B: "Most Americans took a beer."

(C2) A: “Wat namen de meeste Amerikanen?”
B: “De meeste Amerikanen namen een biertje.”

(C3) A: “Namen de meeste Amerikanen een biertje?”
B: “De meeste Amerikanen namen een biertje.”

(C4) Namen de meeste Amerikanen een biertje? ja / nee

7. Drie vrienden gingen gezamenlijk met hun vaders uit eten.
John nam vis. Zijn vader nam vlees.
Michael nam vis. Zijn vader nam ook vis.
Sem nam ook vis. Zijn vader nam de vegetarische schotel.

7. Three friends are taking their fathers out to dinner together.
John took fish. His father took meat.
Michael took fish. His father took fish too.
Sem took fish. His father took the vegetarian dish.

(C1) A: “Hoeveel van de zoons namen vis?”
B: “De meeste zoons namen vis.”
A: “How many of the sons took fish?”
B: “Most sons took fish.”

(C2) A: “Wat namen de meeste zoons?”
B: “De meeste zoons namen vis.”

(C3) A: “Namen de meeste zoons vis?”
B: “De meeste zoons namen vis.”

(C4) Namen de meeste zoons vis? ja / nee

8. Vijf studenten gingen naar een boekwinkel om de boekenbonnen die ze hadden gekregen te verzilveren.
Roderik, een rechtenstudent, koos een roman.
Diederik, een rechtenstudent, koos ook een roman.
Karel, een natuurkundestudent, koos een non-fictie boek.
Harry, een natuurkundestudent, koos ook een non-fictie boek.
Guus, een natuurkundestudent, koos een reisgids.

8. Five students went to the bookstore to cash in the book gift certificates they had.
Roderik, a law student, picked a novel.
Diederik, a law student, also picked a novel.
Karel, a physics student, took a non-fiction book.
Harry, a physics student, also took a non-fiction book.
Guus, a physics student, took a travel guide.

(C1) A: "Hoeveel van de rechtenstudenten kozen een roman?"
B: "De meeste rechtenstudenten kozen een roman."
A: "How many of the law students took a novel?"
B: "Most law students took a novel."

(C2) A: "Wat kozen de meeste rechtenstudenten?"
B: "De meeste rechtenstudenten kozen een roman."

(C3) A: "Kozen de meeste rechtenstudenten een roman?"
B: "De meeste rechtenstudenten kozen een roman."

(C4) Kozen de meeste rechtenstudenten een roman? ja / nee

Exp. 10 Average SI-rates per item

	C1 [wh foc]	C2 [wh n-foc]	C3 [yes/no sent]	C4 [yes/no pol]
1	57	29	14	56
2	14	43	56	29
3	67	29	43	57
4	57	44	43	14
5	67	29	43	57
6	57	44	29	14
7	57	57	29	56
8	29	57	56	43

Exp. 11 Average SI-rates per item

	C1 [wh foc]	C2 [wh n-foc]	C3 [yes/no sent]	C4 [yes/no pol]
1	57	88	20	25
2	38	29	88	40
3	70	25	57	88
4	88	30	25	43
5	70	25	29	25
6	88	50	25	43
7	57	88	40	25
8	38	57	75	20

SAMENVATTING IN HET NEDERLANDS

In de studie van betekenis in natuurlijke taal wordt een onderscheid gemaakt tussen de betekenis van zinnen die te herleiden is naar de betekenis van de woorden en zinsdelen (de 'letterlijke' betekenis), het onderwerp van studie van de *semantiek*, en de betekenis die voortkomt uit de combinatie van de letterlijke betekenis van een zin en de specifieke situatie en context waarin een zin wordt geuit. Deze soort betekenis wordt bestudeerd in de *pragmatiek*. Eén van de meest voorspelbare en wijdverbreide soorten van pragmatische betekenis zijn *Scalaire Implicaturen* (SIs). Dit zijn gevolgtrekkingen gebaseerd op het idee dat bepaalde woorden of woordgroepen een schaal vormen. Intuïtieve voorbeelden hiervan zijn groepen bijvoeglijk naamwoorden. De woorden *vreselijk*, *slecht*, *matig*, *okay*, *goed*, *uitstekend* representeren allemaal een bepaald bereik op een schaal van 'goedheid'. Een scalaire implicatuur is de conclusie van de hoorder dat als een spreker een item van een schaal gebruikt, dezelfde zin met een item dat hoger op de schaal staat onwaar is. Neem bijvoorbeeld de dialoog in (1):

(1) A: "Hoe was het concert?"
B: "Het concert was okay."

Normaal gesproken zal spreker A uit het antwoord van spreker B opmaken dat het concert niet *uitstekend* was. Deze conclusie is een scalaire implicatuur. De betekenis dat het concert niet uitstekend was, is geen onderdeel van de letterlijke betekenis van het woord *okay*, zoals blijkt uit (2):

(2) Het concert was niet okay.

Stel dat de betekenis van *okay* is *okay maar niet uitstekend* (vergelijkbaar met een rapportcijfer tussen 6 en 8), dan zou (2) betekenen dat het concert niet een cijfer tussen de 6 en de 8 scoorde, maar mogelijk dus wel een 9 of een 10. Dat is duidelijk niet wat zin (2) intuïtief betekent, namelijk dat het concert niet tussen de 6 en de 10 scoorde. Daarom blijkt uit (2) dat de letterlijke betekenis van *okay* het bereik 6-10 is, en niet 6-8. Het feit dat de hoorder in (1) normaal gesproken concludeert dat het concert niet uitstekend was (dus niet in het bereik 8-10 viel) is daarom een pragmatische gevolgtrekking.

Deze gevolgtrekking is beschreven door Grice (1967) aan de hand van zijn *Coöperativiteitsprincipe*. Een onderdeel daarvan is dat sprekers in het

algemeen proberen zo informatief mogelijk te zijn gegeven de gesprekssituatie (Grice's *Maxime van Kwantiteit*). Aangezien het rapportcijferbereik 8-10 kleiner is dan het bereik 6-10 is het informatiever om *Het concert was uitstekend* te zeggen dan *Het concert was okay*: de eerste zin maakt een striktere claim over hoe de wereld is. Horn (1972) stelde dat we dit kunnen omschrijven in termen van (asymmetrische) logische *entailment*: in alle gevallen waarin de zin *Het concert was uitstekend* waar is, is de zin *Het concert was okay* ook waar, maar niet andersom. Daarom is de eerste zin informatiever dan de tweede. Gebaseerd op het Coöperativiteitsprincipe gaat een hoorder ervan uit dat een spreker een zo informatief mogelijke zin heeft geuit. Daarom kan de hoorder concluderen dat de spreker de zin met de sterkere scalaire term onwaar acht, anders had zij deze wel verkozen boven de zwakkere zin.

Niet alleen bijvoeglijke naamwoorden brengen scalaire implicaturen teweeg. Er zijn veel verschillende soorten woorden of woordgroepen die een entailment-schaal vormen en dus dergelijke gevolgtrekkingen mogelijk maken. Het voegwoord *of* wordt traditioneel ook beschouwd als een woord dat een scalaire implicatuur kan veroorzaken, gebaseerd op de entailment-relatie met *en*. Vergelijk bijvoorbeeld (3) en (4):

(3) Als Jan een karper of een snoek heeft gevangen, dan eet ik mijn hoed op.

(4) Jan heeft een karper of een snoek gevangen.

De voorwaarde die gesteld wordt in (3) is dat Jan minstens één van beide vissen moet hebben gevangen, en mogelijk allebei. Als hij zowel een karper als een snoek heeft gevangen, moet ik nog steeds mijn hoed opeten. In (3) betekent *of* dus *één van de twee of allebei*, vergelijkbaar met *en/of*. Deze betekenis wordt wel *inclusief-of* genoemd. Maar in (4) is de meest natuurlijke interpretatie dat Jan één van beide vissen heeft gevangen, en niet allebei. Deze betekenis heet *exclusief-of*. Het lijkt er dus op dat *of* ambigu is tussen een inclusieve en een exclusieve betekenis. Echter, Horn (1972) liet zien dat als de basisbetekenis van *of* inclusief is, we de exclusieve betekenis kunnen afleiden met een scalaire implicatuur. Inclusief-*of* vormt een entailment schaal met *en*: In alle gevallen waarin *A en B* waar is, is *A inclusief-of B* ook waar, maar niet andersom. Het is daarom informatiever om *en* te gebruiken dan *of*. Aangezien de spreker dit niet gedaan heeft in (4), moet het wel zo zijn dat de *A en B* situatie onwaar is. Zo komt de exclusieve betekenis van *of* tot stand. In (3) gaat deze redenering niet door, omdat het door de inbedding in een *als...dan* constructie niet informatiever is om *en* te

gebruiken dan *of*. Net als in (2) hierboven, verdwijnt de scalaire implicatuur ook als we *of* inbedden onder *niet*, zie (5):

(5) Jan heeft niet een karper of een snoek gevangen.

Zin (5) betekent dat Jan geen van beide vissen heeft gevangen. De zin is onwaar als Jan zowel een karper als een snoek heeft gevangen. Wat wordt ontkend is dus dat Jan een karper *en/of* een snoek heeft gevangen.

Scalaire implicaturen en in het specifiek de scalaire implicatuur geassocieerd met *of* zijn het onderwerp van deze dissertatie. Het doel van de dissertatie is drieledig. Het eerste doel is te bepalen wat de cruciale eigenschappen van de talige context zijn die bepalen of scalaire implicaturen doorgaan of niet. In de literatuur wordt algemeen onderkend dat het afhangt van de context of een zin met een scalaire term met of zonder een scalaire implicatuur wordt geïnterpreteerd. Toch zijn er nog maar weinig pogingen ondernomen de eigenschappen van de context te vinden die dit bepalen. In deze dissertatie hoop ik bij te dragen aan het vullen van dit gat en zo bij te dragen aan een explicietere theorie van pragmatische betekenis. Het tweede doel komt voort uit de aanpak die is gekozen in dit onderzoek. In de pragmatische literatuur is het gebruikelijk algemene principes te poneren, en deze te illustreren met enkele treffende voorbeelden. Ik stel dat dit niet de juiste manier is om pragmatische theorie explicieter te maken. Pragmatische betekenis is afhankelijk van veel factoren die tegelijkertijd aan het werk zijn (specifieke situatie, spreker, achtergrondkennis etc.) en is onderhavig aan veel individuele variatie, zowel aan de kant van de sprekers als aan de kant van de uitingen zelf. Daarom kunnen we niet op individuele intuïties over specifieke voorbeelden vertrouwen, maar moeten we gecontroleerde experimenten uitvoeren met grotere groepen items en taalgebruikers. Het tweede doel van de dissertatie is dan ook methodologisch van aard: het vinden van een geschikt experimentele methode om pragmatische betekenis, en scalaire implicaturen in het bijzonder, te testen. Het uitvoeren van gecontroleerde experimenten stelt ons bovendien in staat om te testen of de theorieën uit de theoretische pragmatiek goede modellen zijn van het taalsysteem van daadwerkelijke taalgebruikers. Dit is het derde doel van de dissertatie: het testen van de psychologische realiteit van de theorieën van scalaire implicaturen, en specifiek van de scalaire implicatuur-benadering van de betekenis van *of*. De hypothese dat de lexicale betekenis van *of* inclusief is, en hoorders exclusief-*of* afleiden met een scalaire implicatuur, is aantrekkelijk vanuit een theoretisch perspectief, omdat een lexicale ambiguïteit wordt vermeden. Echter, ook hiervoor geldt dat we slechts met gecontroleerde experimenten kunnen achterhalen of deze benadering strookt met de psychologische realiteit.

In hoofdstuk 2 introduceer ik de hypothese die voorkomt uit het werk van Van Kuppevelt (1996) en Van Rooij (2002) dat scalaire implicaturen afhankelijk zijn van de contextuele eigenschap *informatie focus*. Jackendoff (1972) stelde voor dat elke zin gesplitst kan worden in een gedeelte dat nieuwe informatie bevat (de *informatie focus*) en een gedeelte dat oude informatie bevat (de *presuppositie* of *achtergrond*). Van Kuppevelt en Van Rooij voorspellen, vanuit geheel verschillende achterliggende theorieën, dat scalaire implicaturen alleen voorkomen als de scalaire term (bijvoorbeeld *of*) onderdeel is van de informatie focus. Of informatie nieuw of oud is wordt bepaald door de voorgaande context, daarom is informatie focus een contextuele eigenschap. Deze eigenschap kan gemanipuleerd worden door de zin vooraf te laten gaan door een expliciete vraag. Zie bijvoorbeeld (6) en (7):

(6) V: Wat heeft Harry meegenomen?
A: Harry heeft <u>stokbrood of chips</u>$_F$ meegenomen.

(7) V: Wie heeft stokbrood of chips meegenomen?
A: <u>Harry</u>$_F$ heeft stokbrood of chips meegenomen.

Ondanks dat de antwoorden in (6) en (7) dezelfde zin zijn (*Harry heeft stokbrood of chips meegenomen*), zijn hun focus-structuren anders, doordat ze andere vragen beantwoorden. Ten opzichte van de vraag in (6) is *stokbrood of chips* de nieuwe informatie in het antwoord en dus de focus, aangegeven met de onderstreping en het subscript *F*. In (7) is *Harry* de nieuwe informatie en dus de focus. Omdat alleen in (6) de scalaire term *of* onderdeel van de focus is, voorspellen Van Rooij en Van Kuppevelt dat alleen in (6) de scalaire implicatuur (dat Harry niet zowel stokbrood als chips heeft meegenomen) zal worden berekend. Dus in (6) zal *of* eerder als exclusief-*of* worden opgevat, en in (7) als inclusief-*of*. Deze voorspelling test ik in de dissertatie in een serie van 11 experimenten.

In hoofdstuk 3 test ik deze voorspelling in drie experimenten waarin de Truth Value Judgment Task (TVJT) werd gebruikt. In deze experimenten kregen proefpersonen een verhaaltje te lezen waarin *A en B* het geval was. Dit verhaaltje werd gevolgd door een dialoog als in (6) en (7), waarin in het antwoord *A of B* gebruikt werd. Proefpersonen werd gevraagd aan te geven of het antwoord van spreker B waar is. Twee voorbeelditems (uit Experiment 2) zijn (8) en (9):

(8) <u>focus conditie</u>
Julie en Karin waren samen zeedieren aan het zoeken op het strand. Na enig zoeken vond Julie een krab. Niet veel later vond ze ook een zeester. Karin vond helaas helemaal niks. Toen Karin

terugkwam, vroeg haar moeder wat voor zeedieren Julie had gevonden. Karin antwoordde dat Julie een krab of een zeester had gevonden.

Is Karins antwoord waar? waar / onwaar

(9) non-focus conditie
Julie en Karin waren samen zeedieren aan het zoeken op het strand. Na enig zoeken vond Julie een krab. Niet veel later vond ze ook een zeester. Karin vond helaas helemaal niks. Toen ze terugkwamen, vroeg hun moeder wie van hen een krab of een zeester had gevonden. Karin antwoordde dat Julie een krab of een zeester had gevonden.

Is Karins antwoord waar? waar / onwaar

Als een proefpersoon 'onwaar' antwoordt, duidt dit erop dat de proefpersoon *of* als exclusief-*of* heeft geïnterpreteerd en dus de scalaire implicatuur heeft berekend. In Experiment 1 werd naast de expliciete vraag ook het verhaaltje gemanipuleerd om de vraag te ondersteunen. Het resultaat was dat proefpersonen inderdaad vaker 'onwaar' antwoordden in de focus-conditie, wat erop duidt dat de scalaire implicatuur inderdaad vaker werd berekend in deze conditie. In Experiment 2 varieerde alleen de expliciete vraag tussen de condities (zie (8) en (9)). Weer werden significant meer 'onwaar' antwoorden gegeven in de focus-conditie. Dit duidt erop dat informatie focus inderdaad een contextuele eigenschap is die scalaire implicaturen beïnvloedt. In Experiment 3 testte ik of de voorspelling ook uitkomt als de contextuele vraag impliciet is, wat door zowel Van Kuppevelt als Van Rooij wordt gesteld. Om toch de juiste focus-structuur af te dwingen voor de target-zin werden gesproken stimuli gebruikt, waarin het hoofdaccent van de zin werd gevarieerd. Ook werden de verhaaltjes zo opgezet dat ze de impliciete vraag opriepen die overeen kwam met de focus-structuur van de gesproken target-zin. Weer werden er significant meer 'onwaar' antwoorden gegeven in de focus conditie dan in de non-focus conditie, in lijn met de voorspellingen.

In hoofdstuk 4 presenteer ik Experiment 4-7, waarin de voorspelling van focus-gevoeligheid van scalaire implicaturen werd getest met een andere experimentele methode. Het nadeel van de methode gebruikt in Experiment 1-3, de TVJT, is dat de proefpersonen moeten weten wat de echte situatie is. Dit is anders dan hoe scalaire implicaturen in normaal taalgebruik tot stand komen, waar een hoorder juist probeert te achterhalen wat er is gebeurd. Daarom werd in Experiment 4-7 een nieuwe methode gebruikt, genaamd de *Possible World Judgment Task (PWJT)*. Hierin kregen proefpersonen

wederom een verhaaltje en een dialoog te lezen, waarin het antwoord *of* bevatte. Echter, de ware situatie (bijvoorbeeld wie wat had gevonden) was weggelaten, en de proefpersoon werd gevraagd te beoordelen of de *en*-situatie mogelijk was. Een voorbeeld van een item is (10):

(10) Voorbeelditem PWJT
Marieke vertelde aan een vriendin dat Laura en Barbara gisteren zeedieren waren gaan zoeken op het strand, en dat toen ze terugkwamen, hun moeder zei dat als één van hen minstens twee zeedieren had gevonden op het strand, diegene die avond later mocht opblijven.
De vriendin zei: "Oh, en wat had Laura gevonden op het strand?"
Marieke antwoordde: "Laura had een krab of een mossel gevonden op het strand."

Denk je dat het mogelijk is dat Laura zowel een krab als een mossel had gevonden? ja / nee

Een 'nee' antwoord duidt erop dat de proefpersoon de scalaire implicatuur heeft berekend. Zoals in Experiment 1 en 2 werd focus gemanipuleerd door de expliciete vraag in het verhaaltje te variëren. Naast focus testten Experiment 4-7 nog een contextuele eigenschap waarvan geclaimd wordt in de literatuur dat hij belangrijk is voor scalaire implicaturen, door mij 'alternative relevance' genoemd. Dit is de eigenschap dat het sterkere scalaire alternatief (bijv. de zin met *en*) relevant is in de context. Deze eigenschap werd gemanipuleerd door de voorwaarde genoemd in de eerste zin van het verhaaltje te variëren. In (10) is de voorwaarde dat *minstens twee zeedieren* gevangen moesten worden, dus is de zin met *en* (*Laura had een krab en een mossel gevangen*) zeer relevant. In de andere condities was de voorwaarde echter bijvoorbeeld *als een van hen een krab of een mossel had gevonden...* Voor deze voorwaarde is de zin met *en* niet relevant, omdat het niet uitmaakt of ze één van de twee of allebei de dieren hadden gevonden. De resultaten van twee van de vier experimenten lieten een significant verschil zien in 'nee'-antwoorden tussen de conditie met focus en alternative relevance en de conditie zonder beide eigenschappen. Dit duidt erop dat beide eigenschappen samen van belang zijn voor scalaire implicaturen. De verschillen tussen de condities die verschilden op één eigenschap waren te klein om uitspraken te kunnen doen over hun onderlinge verhouding. Ik beargumenteer dat het feit dat de verschillen kleiner zijn dan in Experiment 1-3 komt doordat de experimentele vraag de kritieke manipulaties overschaduwt. Omdat deze vraag expliciet naar de kritieke gevolgtrekking vraagt, maakt hij deze relevant in alle condities. Daarom stel ik dat we een

experimentele methode moeten gebruiken waarbij we de echte situatie kunnen weglaten (versus de TVJT), en tevens de experimentele vraag kunnen weglaten (versus de PWJT). De manier om dit te bereiken is door taalverwerkingsexperimenten uit te voeren.

In hoofdstuk 5 bespreek ik taalverwerkingsexperimenten uit de literatuur waarin naar scalaire implicaturen werd gekeken. Een aantal verwerkingsexperimenten laat een toegenomen verwerkingslast zien die mogelijk het gevolg is van het berekenen van een implicatuur. Ik besteed specifiek aandacht aan de experimenten van Katsos (2006), die laat zien dat zinsdelen met scalaire termen langzamer gelezen worden in contexten die scalaire implicaturen veroorzaken. Deze waarnemingen neem ik als uitgangspunt voor Experiment 8 en 9, gepresenteerd in hoofdstuk 6. In deze experimenten werd de Self-Paced Reading methode gebruikt. In deze methode lezen proefpersonen een tekst op een scherm stukje voor stukje (woord of woordgroep), en bepalen ze zelf wanneer het volgende stukje zichtbaar wordt door op een knop te drukken. Op die manier kan de tijd besteed aan een stukje worden vergeleken tussen condities. Als het berekenen van een scalaire implicatuur extra verwerkingskosten met zich meebrengt, zou de scalaire regio langzamer moeten worden gelezen in de focus conditie dan de non-focus conditie. In Experiment 8 werden hetzelfde type verhaaltjes gebruikt als in (TVJT) Experiment 2 (zie (8) en (9) hierboven). Een repetitie-effect dat mogelijk de leestijden zou beïnvloeden werd gecontroleerd d.m.v. twee controle-condities met *en* i.p.v. *of*. Er werd echter geen vertraging geobserveerd in de focus conditie vergeleken met de non-focus conditie. Ik stelde dat dit mogelijk het gevolg is van het minimale karakter van de verhaaltjes. Omdat er geen gevolgen voor het verhaaltje waren van het gebruik van *of* i.p.v. *en*, en er nooit naar gevraagd werd in de controlevragen, was er mogelijk geen reden voor de proefpersonen om de implicatuur te berekenen. Daarom werd in Experiment 9 gebruik gemaakt van de rijkere verhaaltjes van (PWJT) Experiment 5 (zie bijvoorbeeld (10) hierboven). Door de voorwaarde die was opgenomen in deze verhaaltjes (bijv. het vinden van tenminste twee zeedieren in (10)) was de scalaire implicatuur belangrijk voor de uitkomst van het verhaaltje. Dit gaf proefpersonen een reden om de implicatuur te berekenen. Als gevolg van het gebruik van deze verhaaltjes werd het cumulatieve effect van focus en alternative relevance getest. Experiment 9 bevatte ook een extra meetpunt. De target-zin werd gevolgd door een zin waarin het meervoud persoonlijk voornaamwoord *ze* werd gebruikt om terug te verwijzen naar *A of B*. Als de proefpersoon de scalaire implicatuur heeft berekend voor de target-zin en dus heeft geconcludeerd dat *A en B* niet het geval is, dan is dit meervoud onverwacht. In dat geval zou het tot een vertraging moeten leiden vergeleken met de controleconditie met het enkelvoudig persoonlijk voornaamwoord:

'm. De resultaten lieten zien dat zowel de scalaire regio zelf als *ze* in de volgende zin significant trager gelezen werden in de focus-conditie dan in de non-focus conditie. Dit verschil werd niet geobserveerd voor de controlecondities met *en* i.p.v. *of* en de controlecondities met *'m* i.p.v. *ze*. Dit is een sterke aanwijzing dat er inderdaad meer scalaire implicaturen werden berekend in de focus-conditie. Daarnaast suggereert het dat het berekenen van een scalaire implicatuur extra verwerkingskosten met zich meebrengt, wat een belangrijke observatie is in het debat over hoe scalaire implicaturen tot stand komen.

In hoofdstuk 7 presenteer ik Experiment 10 en 11, waarin ik controleer of de voorspelling dat scalaire implicaturen gevoelig zijn voor focus ook uitkomt voor een andere scalaire term in een andere positie. Ik gebruik in deze experimenten de scalaire term *de meeste* in subject positie. De focus-structuur werd wederom gemanipuleerd d.m.v. expliciete vragen, bijvoorbeeld (11) en (12):

(11) V: Hoeveel van de studenten dronken er bier?
A: De meeste$_F$ studenten dronken bier.

(12) V: Wat dronken de meeste studenten?
A: De meeste studenten dronken bier$_F$.

De zin *De meeste studenten dronken bier* kan de scalaire implicatuur triggeren dat *niet alle* studenten bier dronken, gebaseerd op de entailment schaal <meeste, alle>. De voorspelling is dat deze implicatuur vaker zal worden getrokken in (11) dan in (12), aangezien alleen in (11) *de meeste* onderdeel is van de focus. Experiment 10 en 11 maakten weer gebruik van de TVJT. Na een verhaaltje waarin alle studenten bier dronken, werd het antwoord in (11) inderdaad significant vaker als 'onwaar' beoordeeld. Het effect van focus-gevoeligheid werd dus gerepliceerd voor een andere scalaire term in een andere positie. Daarnaast bevatten Experiment 10 en 11 een conditie waarin dezelfde zin een antwoord was op een ja/nee-vraag, bijv. *Dronken de meeste studenten bier?* Het percentage implicaturen in deze conditie was vergelijkbaar met de non-focus conditie. Dit is in lijn met de focus-gevoelige theorieën omdat de target-zin in deze conditie geen nieuwe informatie bevat en dus geen focus-gedeelte heeft.

In hoofdstuk 8 ga ik verder in op de psychologische realiteit van de scalaire implicatuur-benadering van *of*. Een bekend probleem van deze benadering is dat een zin als *A of B* ook een andere implicatuur teweeg brengt: de gevolgtrekking dat de spreker niet weet of A waar is, en niet weet of B waar is (vaak *Clausal* of *Ignorance* Implicatures genoemd). Dit is problematisch voor de traditionele Griceaanse benadering, waarin de hoorder

moet veronderstellen dat de spreker geïnformeerd is over de waarheid van de *A en B* zin, anders kan de hoorder niet concluderen dat deze onwaar is. Ik noem dit probleem de *Speaker Expertise Paradox*. In hoofdstuk 8 presenteer ik additionele data van Experiment 4, waaruit blijkt dat proefpersonen inderdaad zeer frequent scalaire implicaturen berekenen terwijl ze de spreker niet als compleet geïnformeerd beschouwen. In het tweede deel van het hoofdstuk presenteer ik vier alternatieve theorieën uit de literatuur die dit probleem mogelijk omzeilen. Ik bespreek de twee Griceaanse alternatieven van Sauerland (2004) en Van Rooij & Schulz (2004), die beide een complexer mechanisme voorstellen om de aanname van de hoorder over de kennis van de spreker te modelleren. Daarna bespreek ik twee theorieën die het idee verlaten dat scalaire implicaturen pragmatische gevolgtrekkingen zijn en die beweren dat scalaire implicaturen worden berekend in het syntactisch-semantische systeem. Dit zijn de theorieën van Fox (2007) en Chierchia (2004/2006). Ik beargumenteer dat alle vier de theorieën in hun huidige vorm niet zowel de focus-data kunnen verklaren als een oplossing bieden voor de Speaker Expertise Paradox.

In hoofdstuk 9 stel ik de implicatuur benadering van *of* verder op de proef door de relatie tussen scalaire implicaturen en exhaustiviteit te bestuderen. Exhaustiviteit is de gevolgtrekking dat het antwoord in (13) meestal wordt opgevat als (14):

(13) V: Wie heeft na de les zijn excuses aangeboden?
A: Bart heeft zijn excuses aangeboden.

(14) Alleen Bart (van de relevante set mensen) heeft zijn excuses aangeboden.

Volgens de theorie van Van Rooij (2002), die de focus-gevoeligheid van scalaire implicaturen voorspelde, zijn scalaire implicaturen en exhaustiviteit beide het gevolg van de toepassing van dezelfde exhaustiviteits-operator. Dit brengt een interessante voorspelling met zich mee: voor een zin als (15) zou de scalaire implicatuur in (16) altijd gepaard moeten gaan met de exhaustiviteitsinferentie in (17), omdat beiden resultaat zijn van hetzelfde proces.

(15) Bart heeft een spin of een slang gevangen.

(16) Bart heeft niet een spin en een slang gevangen.

(17) Bart heeft niet nog iets anders gevangen (naast een spin of een slang).

Ik beargumenteer in hoofdstuk 9 dat niet alleen de theorie van Van Rooij (2002) (en Van Rooij & Schulz 2004) deze voorspelling maakt, maar ook de grammaticale theorie van Chierchia en collega's. Experiment 5-7 bevatten naast items die scalaire implicaturen testten ook items die de 'externe' exhaustiviteits-inferenties als in (17) testten. Daaruit blijkt dat de twee gevolgtrekkingen niet altijd samen voorkomen: in twee van de drie condities in Experiment 5 en één van de drie condities in Experiment 6 en 7 werden significant meer scalaire implicaturen geobserveerd dan externe exhaustiviteits-inferenties. Dit wekt twijfel over de claim dat beide het gevolg zijn van één mechanisme, en over de psychologische realiteit van de implicatuur-benadering van de betekenis van *of*.

Hoofdstuk 10 bevat een samenvatting van de dissertatie. In dit hoofdstuk presenteer ik tevens de conclusies en de suggesties voor verder onderzoek voor de drie doelen van de dissertatie. Het eerste doel was onderzoeken welke eigenschappen van de context bepalend zijn voor de aanwezigheid van scalaire implicaturen. Ik concludeer dat het effect van de contextuele eigenschap informatie focus in een wijd bereik van experimentele condities is bevestigd, en er aanwijzingen zijn dat ook alternative relevance van belang is voor scalaire implicaturen. Wat betreft het tweede, methodologische doel is gebleken dat een impliciete taak zoals in de verwerkingsexperimenten de beste methode is om de aanwezigheid of afwezigheid van scalaire implicaturen te onderzoeken, aangezien hier zowel de echte situatie als een expliciete experimentele vraag weggelaten kunnen worden. Het derde doel was het testen van de psychologische realiteit van de scalaire implicatuur-benadering van de betekenis van *of*. Dit heb ik gedaan door data te verzamelen over het verband tussen de interpretatie van *of* en vier zaken: het effect van focus, de verwerkingskosten, de kennis-staat van de spreker en exhaustiviteit-inferenties. In de loop van de dissertatie heb ik vier soorten implicatuur-theorieën besproken, die samen een groot deel van de dominante theorieën in de literatuur omvatten: de Griceaanse/PCI-benadering, de default benadering, de exhaustiviteitsbenadering van Van Rooij & Schulz (2004) en de grammaticale benadering van Chierchia en collega's. Ik beargumenteer in hoofdstuk 10 dat geen van deze vier benaderingen alle vier typen data die zijn gepresenteerd in deze dissertatie kan verklaren. Dit wekt twijfel over de psychologische realiteit van de scalaire implicatuur benadering van de inclusief/exclusief-*of* ambiguïteit. Ik eindig daarom de dissertatie met de suggestie voor verder onderzoek naar de hypothese dat exclusief-*of* de basale betekenis is en inclusief-*of* hiervan afgeleid is door middel van een existentiële kwantificatie.

CURRICULUM VITAE

Arjen Zondervan was born on August 28, 1980 in Amsterdam. After obtaining his gymnasium diploma at the Amstelveen College in 1998, he studied philosophy at the University of Amsterdam and obtained his propeadeutic (first year) degree in 1999 (cum laude). In 1999-2000 he studied architecture at Delft University of Technology. In 2000 he started his studies in linguistics at Utrecht University. He obtained his bachelor's degree in linguistics in 2004 (cum laude), and joined the international MPhil program in linguistics of UiL-OTS. As a part of this program he was a visiting graduate student at the Linguistics Department of the University of Massachusetts at Amherst in the fall semester of 2005. He completed the tracks semantics and experimental psycholinguistics and obtained his master's degree in 2006 (cum laude) with a thesis on scalar implicatures. In September of that year he started his AiO (PhD) project on contextual effects on scalar implicatures, of which this dissertation is the result.