On the pragmatics of numeral modifiers

The availability and time course of variation, ignorance and indifference inferences

Published by

LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111
e-mail: lot@uu.nl
http://www.lotschool.nl

Cover illustration: LingLab. By Maria Koutsodimitropoulou; edited by Kalliopi Mira.

ISBN: 978-94-6093-293-9
NUR: 616

On the pragmatics of numeral modifiers

The availability and time course of variation, ignorance and indifference inferences

Over de pragmatiek van telwoord-modificeerders

De beschikbaarheid en het tijdsverloop van gevolgtrekkingen van variatie, onwetendheid en onverschilligheid

(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. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
maandag 2 juli 2018
des middags te 12.45 uur

door

Stavroula Alexandropoulou

geboren 25 april 1986
te Athene, Griekenland

Promotor:	Prof. dr. H.E. de Swart
Co-promotor:	Dr. R.W.F. Nouwen

The research reported here has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement no. 313502.

Στον θείο τον Σωτήρη

Contents

Acknowledgements

We're on the road to nowhere—just started playing on the radio—but it's high time to make a stop and note all these people that had a hand in making my PhD trip smoother, more inspiring and fun, and complete, I guess.

I'm very grateful to my supervisor Rick Nouwen, to begin with, that made it real for me. After all these years, all these meetings, after exchanging experiences with fellow students, I can safely say that you've been a true advisor, and unique, Rick. Thank you for all the knowledge you've given me and for that you led me to find out myself as a real master of the socratic method, for your patience, your supportiveness and actual help throughout the various stages of my project, for sharing your genius-ness with me. I'm not sure how far I would have gotten on my own without your influence and inspiration, and that's partly why "what's next?" is a question that makes me feel uncomfortable.

I'm also grateful to my second supervisor, Henriëtte de Swart, who although I met less regularly, my meetings with her had a significant impact on me and my project; I would always leave having learned something new. Henriëtte, thank you for teaching me how to always seek the bigger picture of my research through your tactfully asked (tough) questions, for your support and trust on me, for reminding me to be confident, also in my way of writing. I was so amazed to realize that you literally scrutinized my manuscript, giving those crucial comments and advice that made it into a book.

The other ROSE project members: Dominique Blok, Yaron McNabb, and Sofia Bimpikou. Thank you, guys, for our joint work, meetings, reading groups, discussions, dinners, lunches, coffees, drinks. It was a pleasure!

... and Jakub Dotlačil of course, who feels to me like he has always been part of ROSE, and in a sense my third—brilliant—supervisor. Jakub, you've taught me so much. Thank you for introducing me to eye-tracking and to more advanced experimental research, for patiently listening to and answering my full of despair and agony questions. I so much enjoyed our collaboration! Your contribution to this book is just enormous.

I would further like to thank the members of my reading committee for taking the time to read my manuscript: Chris Cummins, Petra Hendriks, Petra Schumacher, Jos van Berkum, and Yoad Winter.

My two paranymphs, Elli and Myrto, and my two substitute paranymphs, Maartje and Sophia. Ellaki and Myrtoulaki, I feel so lucky to have met you in that glorious May of 2006 at the occupied buildings of the faculty of philosophy in Thessaloniki and to be together ever since, in our university years, in our MA adventure here in Utrecht, but also afterwards, from distance. Thank you for aaaaaaaall your unconditional, never-ending support and help, especially during the months of my dissertating, and apologies for overloading you with my questions till the very last moment. Maart(jetje) and Soph (mein anderes ich ;-), thank you for being there all these years and for assisting me in so many different ways; your help with my experiments and stats was honestly invaluable.

There were many more people that had a substantial contribution to the different stages of my experiments. Thank you, Carolien van den Hazelkamp, Marko Simonovic, Sylvia Wijsman, Charlotte Lindenbergh, Ilse Angenent, Sofia Bimpikou, Ying Liu, Heidi Klockmann for helping with the construction of items and cleaning eye tracking data; Onno Kosters, Ileana Grama, Annemarie Kerkhoff, Sergey Avrutin, Shalom Zuckerman, Willemijn Heeren, Yvette Linders for letting me distribute my questionnaires to your students; Arnout Koornneef, Iris Mulders, Maarten Duijndam, Martijn van der Klis, Chris van Run for all your feedback and assistance while running my experiments at the UiL-OTS lab; Mattis van den Bergh, Kirsten Schutter, Piet van Tuijl, Björn 't Hart, Yipu Wei for discussing stats with me, and of course all my participants!

Special thanks to Alexis Dimitriadis for creating the Lotdiss.cls and making our lives during writing much easier, to those people that did not turn down my last minute request and accepted to translate a long summary into Dutch (Samenvatting in het Nederlands): Rick Nouwen, Björn 't Hart, Eva Poortman, Anja Goldschmidt, and Yvette Linders, and lastly to my sister-friends (in a different sense each one) Maria Koutsodimitropoulou and Kalliopi Mira for our brainstormings over the front cover, and for materializing our thoughts and my vague desires.

During my PhD years I had the opportunity to attend a variety of conferences, workshops, reading groups, seminars or informal meetings, where I met and talked to people, whose input was useful for my research. For that I would like to thank Lisa Bylinina, Liz Coppock, Maria Aloni, Jacopo Romoli, Maria Spychalska, Teodora Mihoc, Brian Buccola, Jon Ander Mendia, Andreea Nicolae, Anastasia Giannakidou, Alexandre Cremers, Bernhard Schwarz, Floris Roelofsen, Francesca Delogu, Maria Barouni, Winfried Lechner, and possibly a few others that I'm now forgetting. I have also benefited a lot from my participation in the Experimental Linguistics Talks in Utrecht (ELiTU), the Leiden Utrecht Semantics Happenings (LUSH), the Weak Referentiality talks, and the ROSE talks.

Being part of UiL-OTS since 2011, over the years, I started feeling like

home at Trans, and this is thanks to everyone having been and being around: my colleagues, the staff, my officemates. I'd like to thank each and everyone of you, at the attic, at the kitchen with the colorful chairs, at the Cohen zaal, from the kibbutz, from the polish water parties, from the Friday drinks, from the Chinese dinners, from the Sunday mornings with pancakes, the picknicks at Lepelenburg, the garden lunches, the Dutch classes with dinner, the Thanksgivings, the Uiltjesdagen activities, from the lab, from my various offices and in particular from my latest one: Roos, Sanne, and Yvette, who like my wiser sisters never ceased to coach me.

Outside the UiL-OTS world, there are a few people that really made my last five years in Utrecht carefree and meaningful. Thank you, Katerina, for your kindness and for respecting my request to not call me after work these last months, unless I call you (eh, I couldn't resist to a coffee at the Village), for the nice, welcoming shelters you've created with Alex; thank you both for your support, for our endless discussions and smokes, also with Ilias and Eleni, for our linguistic brainstormings (that one with Nikos and Dimitris seriously helped me) as well as for our heated arguments, Aleko—you and Nikos are soooo WRONG about awareness! *Κουράγιο και τελειώνουν τα παλούκια, μανάκι.*

My deepest thanks to Elena (Servokypria) for motivating me to get up, go to the library and spend all these hours of inspiration and coffee there in my last months of writing.

Last but not least, I want to thank my long-distance people: my family-friends and my extended family. Your infinite support, love, and patience mean a lot to me. Thank you for showing me large amounts of them during this last year.

CHAPTER 1

Introduction

1.1 The meaning of modified numerals

Languages have a quite large inventory of expressions to refer to an imprecise quantity such as $n \geq 4$. In English, for instance, this quantity can be described by the expressions *4 or more*, *at least 4*, *more than 3*, *minimally 4*, *above 3*, *over 3*, etc. Crucially, this abundance of expressions, called modified numerals, is not a quirk of English, but rather a cross-linguistic phenomenon (see Nouwen, 2010). The first question that comes to mind then is *why does language provide us with so many different tools in order to express the exact same quantity?* Do all these expressions actually convey the same meaning? The simple answer is no. If anything, modified numerals became known in the semantics/pragmatics literature because only some of the truth-conditionally equivalent expressions above, e.g., *at least 4*, *4 or more*, *minimally 4*, can additionally convey a signal that the speaker is ignorant about the exact quantity n. That is, if I tell you (1), you will probably infer that I happen to lack the information as to the exact number of souvlakis Magda had.

(1) Magda ate at least four souvlakis.

There is consensus now that that these expressions differ with respect to these so-called *speaker ignorance* effects, for example, and that these effects are pragmatic in nature. In other words, while their core quantitative meaning is the same, as suggested in the beginning, there is a difference with respect to additional meanings that may be triggered. As such modified numerals are a key phenomenon for the study of the semantics-pragmatics distinction. The present

thesis will investigate the speaker ignorance effects as well as two other types of meaning, i.e., *variation* effects and *speaker indifference* effects, by looking at the off-line and the real-time comprehension of utterances with two distinct kinds of numeral modifiers. The overarching goal is to demarcate the core meaning of modified numerals from their pragmatic meaning by defining the latter in detail.

1.2 The semantics-pragmatics divide: The theoretical and experimental landscape

The debate as to the precise division of labor between semantics and pragmatics is particularly popular in the theoretical and the experimental research on conversational implicatures. Let us start with the relevant theoretical landscape and naturally move to the experimental one.

A conversational implicature is a defeasible non-conventional inference that is, for instance, derived by the hearer on the assumption that the speaker observes the conversational maxims (Quality, Quantity, Manner, and Relation) or the Cooperative Principle (Grice, 1975). Grice has made a distinction between two types of conversational implicatures: the generalized and the particularized ones. While the latter are contingent on special features of the general context, the former are *normally carried* by a certain expression (Grice, 1975, p. 56). The well-known scalar implicatures associated with the scalar quantifier *some* are an instance of generalized conversational implicatures, see example below.

(2) Some implicatures are conversational.
$\rightsquigarrow$ *not all implicatures are conversational*

Three of the most familiar or widely known theoretical approaches to scalar implicatures are the Gricean view, the grammatical view, and Levinson's (2000) default view on implicatures. These approaches diverge on the distinction they establish between the semantic and the pragmatic meaning, and on how tightly these two are taken to be related. On the Gricean view, the scalar implicature is computed post-compositionally, after the literal/semantic meaning is computed; under Levinson's account, it is encoded in the lexical content of the scalar expression and it is a default inference, thus, rapidly derived, while, quite similarly, on the grammatical view, it is derived, locally, in the grammar via a regular compositional process (Chierchia, 2006; Chierchia, Fox & Spector, 2013).

Of the three approaches only Levinson's is also a psychological theory and makes claims about processing, namely, that scalar implicatures are derived early in on-line comprehension, independently of context. However, as Geurts (2011) points out, supplementing the Gricean theory with the assumption *if [...] ϕ implicates that ψ, and hearer H actually infers ψ from ϕ, then H has derived ψ as an implicature* (p. 66) can make it into a psychological theory and

can *yield an ample supply of predictions that can be tested in the lab* (p. 67). This is a necessary assumption for testing the availability of implicatures, for instance, in simple off-line judgement tasks (cf., e.g., the truth-value judgement task on scalar implicatures in Bott & Noveck, 2004). That is, such an assumption constitutes a necessary condition for crossing from theoretical pragmatics to experimental pragmatics. Still, though, making this extra assumption that the listener will derive the implicature intended by the speaker does not specify how and under which circumstances listeners compute an implicature, and whether this happens across different contexts. The largest part of experimental pragmatics is actually concerned with how a listener computes an implicature during incremental interpretation of an utterance in a certain context. Given that theoretical accounts do not include any relevant prediction, with the exception of Levinson (2000), a number of psychological theories have been developed for the computation of scalar implicatures. Those theories make well-defined predictions as to the time course of scalar implicatures and also differ as to how they approach the semantics/pragmatics divide.

The two-level language processing model put forth by Huang & Snedeker (2009) is such a theory. According to this theory, semantic information, or the literal interpretation, is processed first, and pragmatic or higher-level processes become available later in incremental interpretation. Thus, the prediction for the computation of scalar implicatures is that first the lower-bound semantic interpretation (*at least some*) is computed, rapidly, independently of context, while scalar inferences are delayed and extra time is needed for their computation.

There is also a psycholinguistic view of language interpretation that takes context into account, the so-called Context-Driven view. Context-Driven accounts claim that the scalar implicature associated with a scalar term needs contextual support in order to be calculated on-line, while in a neutral/ underspecified context only the lower-bound semantic interpretation is available, which is the basic interpretation of the scalar term (see Breheny, Katsos & Williams, 2006; Katsos & Cummins, 2010, and references therein). That is, contrary to Levinson's (2000) account where the scalar inference is always calculated by default and effortlessly, on Context-Driven accounts a scalar implicature is calculated only if there are contextual cues, if the strengthened interpretation is relevant in the context, and this calculation requires extra time and processing resources.

Finally, there is the so-called Constraint-Based approach to scalar implicatures proposed by Degen & Tanenhaus (2014). This is in fact a Context-Driven approach. However, it specifically posits that *as an utterance unfolds, listeners rapidly integrate multiple sources of information. That is, utterance comprehension is probabilistic and constraint-based* (p. 671) and that the more probabilistic support from contextual cues (e.g., cf. the partitive *summa* vs. *some*, the availability and naturalness of lexical alternatives to *some*), the faster and easier a scalar inference is computed. On the other hand, the computation of an inference incurs a processing cost when there is not enough support of this

kind. Given that, Degen & Tanenhaus claim that the semantic/pragmatics distinction does not map onto a distinction in real-time processing.

The existing findings as to the processing of scalar implicatures are conflicting (see Chemla & Singh, 2014a,b; Sauerland & Schumacher, 2016 for a review of relevant experimental studies), creating a lively debate as to whether the derivation of such implicatures incurs a processing cost (see, e.g., Bott, Bailey & Grodner, 2012; Bott & Noveck, 2004; Breheny et al., 2006; Huang & Snedeker, 2009; Grodner, Klein, Carbary & Tanenhaus, 2010; Tomlinson, Bailey & Bott, 2013 *vs.* Breheny, Ferguson & Katsos, 2013; Foppolo & Marelli, 2017; Grodner et al., 2010, among others). However, it is established by now that that their computation is context-sensitive. This means that even when contextual considerations are taken into account in the experimental design, there are experiments that provide evidence that the computation of scalar implicatures is effortful (e.g., Breheny et al., 2006) and others that demonstrate that it is rapidly performed without any extra processing cost (e.g., Degen & Tanenhaus, 2014; Foppolo & Marelli, 2017). There is, moreover, a growing neurolinguistic literature on the computation of scalar implicatures, delivering divergent results, too, as far as their time course is concerned, also when compared to the corresponding semantic interpretations (see, e.g., Chevallier, Bonnefond, van der Henst & Noveck, 2010; Nieuwland, Ditman & Kuperberg, 2010; Noveck & Posada, 2003; Politzer-Ahles, Fiorentino, Jiang & Zhou, 2012; Politzer-Ahles & Gwilliams, 2015; Spychalska, Kontinen & Werning, 2016). Hence, so far as the cognitive signature of implicature calculation is indeterminate, we are still in search of the right and suitable psycholinguistic hypotheses to complement the pragmatic theory and thus more research and evidence are in order. In this light, I view the present thesis as an attempt to provide new and valuable insights to the ongoing debate under discussion. Let us see how studying modified numerals can benefit the psycholinguistic and the theoretical pragmatics research.

1.3 Modified numerals: A case study in experimental semantics-pragmatics

The domain of modified numerals is rich in inferences, predominantly argued not to be part of, or entailed by, the core meaning of numeral modifiers, but rather to be derived via some pragmatic mechanism. Surprisingly enough, modified numerals have received significantly less attention as regards the experimental investigation of the implications they are associated with, as compared to other quantity expressions. Consider, again, the scalar quantifier *some*, perhaps the most salient case of quantity expressions in experimental pragmatics, as it has extensively been studied with respect to the scalar implicature it triggers and its processing (see relevant references in previous section). On a theoretical level, I believe that looking at more, comparable phenomena can

provide a better and more concrete picture of the boundary between the standard grammatical or compositional computations and the pragmatic principles that lead listeners to draw the relevant inferences. Additionally, it can possibly offer insightful information regarding the growing realization that there is considerable diversity in drawing implicatures associated with quantity expressions (see, for instance, Doran, Baker, McNabb, Larson & Ward, 2009; Doran, Ward, Larson, McNabb & Baker, 2012; van Tiel, van Miltenburg, Zevakhina & Geurts, 2016 on scalar implicatures). Moreover, on a psycholinguistic level, looking into how other pragmatic inferences become available and are processed could contribute further insights into the vague area between pragmatic theory and language processing, and help bridge the theoretical and the experimental facets of pragmatics summarized in the previous section. Specifically, such an attempt can improve our understanding of how pragmatic principles govern and modulate the interpretation of quantity expressions, but also how and to what extent those principles manifest themselves in real-time language processing.

In this thesis, I will combine theory and experiment in a way that they will inform and benefit from one another. I will consider different theoretical accounts of modified numerals, and extract and assess predictions they make in relation to the research questions each study sets out to answer focusing on three types of implication of modified numerals. The relevant predictions are not always given explicitly by theorists, or sometimes they are not even transparent or existing, especially as far as the processing of modified numerals is concerned. When predictions are not straightforward, I will rely on related existing uncontroversial empirical evidence to formulate them. However, it will be hard to do the same in the case of the processing of modified numerals and their implications. On the one hand, almost none of the theoretical proposals touch upon the processing of numeral modifiers and of their (pragmatic) implications; on the other hand, no specific predictions can be formulated given evidence on the processing of implicatures associated with quantity expressions in general, as the existing evidence is conflicting, albeit plentiful. Hence, given also that more generally the theoretical frameworks that the various modified numeral proposals are couched in (e.g., (neo-)Gricean, grammatical, inquisitive, and speech-act approaches) do not incorporate any processing considerations about pragmatic inferences (for relevant discussion re Gricean pragmatics see, e.g., Bach, 2006; Geurts & Rubio-Fernández, 2015, but see also Noveck & Reboul, 2008), I will not take sides as to whether the derivation of a pragmatic inference by the listener is an effortful process or not. Thus, I will consider both processing possibilities when dealing with the relevant predictions.

1.4 Research goal and questions

The present thesis will probe the pragmatics of modified numerals in a series of experiments. More specifically, it seeks to obtain insights into the availability and likelihood of three different inferences—variation, speaker ignorance,

speaker indifference—their semantic/pragmatic nature, their strength and the underlying mechanism of derivation, ultimately drawing a parallel among the three of them as well as a comparison between numeral modifiers. This study will help adjudicate among the different existing theoretical accounts of numeral modifiers and will specifically demonstrate which components of those accounts can capture the observed data. Importantly, although the majority of the available accounts adopt a pragmatic take on the speaker ignorance or variation effects, for instance, the present study will further allow us to distinguish even between these pragmatic accounts. Drawing and building on some of them, it will eventually point to what a comprehensive theory of numeral modifiers should look like in order to accommodate the overall findings as to the three different inferences under investigation in relation to each kind of numeral modifier tested. The individual questions that will be addressed and will guide the thesis towards achieving the general goal are grouped as follows:

- ▷ What is the likelihood of drawing a variation inference? Do different (truth-conditionally equivalent) modified numerals give rise to variation effects to the same extent? What is the semantic/pragmatic status and strength of variation effects? What type of variation do modified numerals involve and what are the implications for their underlying mechanism of derivation?

- ▷ How are ignorance inferences accessed in incremental interpretation of modified numerals? What are the insights into the nature and strength of speaker ignorance inferences and into the underlying mechanism of derivation? How is this mechanism different across different modified numerals? How does the mechanism responsible for speaker ignorance effects relate to that responsible for variation effects?

- ▷ Is speaker indifference an available implication of modified numerals? How could we capture speaker indifference and how does it relate to speaker ignorance? Does it manifest itself in incremental interpretation of modified numerals, also in comparison with speaker ignorance effects?

1.5 Structure of the thesis

The present thesis is organized into eight chapters. The current introductory chapter will be followed by **chapter 2**, where I will review the various existing theoretical accounts of speaker ignorance and variation effects, highlighting certain aspects that will be crucial for the experiments to come: i.e., source of alternatives fed into the mechanism of derivation, obligatoriness of drawing the relevant inferences, and specification of alternatives.

Chapter 3 explores variation effects of different modified numerals by providing novel experimental data from five experiments using a felicity judgement task (variation experiments 1a-c and 2a-b). I show that partial variation effects

are available with superlative, disjunctive as well as with comparative modified numerals, and that they are non-obligatory pragmatic implications. I further show that these variation effects are more specifically of the so-called specific variation type, where the partial variation reading necessarily includes an inference as to the minimum value compatible with the modified numeral. This has implications for the type of alternatives fed into the pragmatic mechanism associated with modified numerals, excluding theoretical proposals that only consider *exact* alternatives. Lastly, based on indications stemming from our experimental investigation, I argue that variation effects become less robustly available with comparative modifiers, pointing to a different pragmatic mechanism of derivation, whereby alternatives come from a different source as compared to those of superlative modifiers.

Chapter 4 reviews the existing experimental studies on speaker ignorance effects. I discuss problematic issues of the indirect investigations of speaker ignorance as well as of the few more direct ones. I argue that, although these studies do yield substantial results, namely, that speaker ignorance effects are pragmatic in nature and are available with both superlative and comparative modifiers, still more needs to be explored so as to better understand the pragmatic mechanism responsible for these effects and to what extent this is the same in superlative and comparative modifiers. These issues are not tackled by the existing experimental literature and are, thus, taken to be the starting point of the experiments on speaker ignorance reported in this thesis.

Chapter 5 presents an eye-tracking reading experiment that directly probes ignorance effects of the superlative numeral modifier *at least* and the time course thereof, aiming to obtain insight into how and when ignorance effects arise as well as into the nature of the mechanism that is responsible for their derivation. A processing penalty is found to be associated with the modified numeral phrase in the context that biases a speaker ignorance interpretation. This result is taken to be due to the on-line calculation of a speaker ignorance implicature. The overall findings disfavor a number of theoretical accounts that derive an obligatory context-independent type of speaker ignorance with superlative modifiers. Furthermore, they are taken to be in favor of neo-Gricean accounts of superlative modifiers that derive context-sensitive specific speaker ignorance implications via a Quantity-based reasoning, where the ignorance reading includes an inference as to the minimum value compatible with the modified numeral. Lastly, I point out a couple of possible confounds challenging the relevant finding, which I address and set out to eliminate in the second eye-tracking experiment.

Aiming to solidify the findings of the first eye-tracking experiment, I ran a follow-up eye-tracking experiment, which I report in **chapter 6**. By probing speaker ignorance with comparative modified numerals too and modifying slightly the material, this experiment does away with the confounds under consideration and replicates the exact same finding as the first eye-tracking experiment. Specifically, this experiment brings to light—on-line or indirect—traces of the derivation of specific speaker ignorance inferences. Thus, it is

firmly concluded that superlative modifiers trigger a specific speaker ignorance implication that is context-depended and pragmatic in nature, as captured by neo-Gricean Quantity-based accounts. Moreover, this experiment reveals a difference in the processing profiles of superlative and comparative modifiers, with the processing of the latter modifier condition suggesting that specific speaker ignorance is available but more loosely associated with comparative modified numerals as compared to superlative ones. Finally, based on the results of the two eye-tracking experiments on speaker ignorance and on the results of the variation experiments, I put forth a composite theory that captures speaker ignorance and partial variation effects via the same pragmatic mechanism for both superlative and comparative modifiers. On this proposal, the two numeral modifiers are taken to be associated with the same type of alternatives, though part of them (i.e., *exact* alternatives) is motivated differently for the two numeral modifiers. This difference in the source of the specific alternatives is taken to be responsible for the strength difference attested between superlative and comparative modifiers as to specific variation and specific speaker ignorance effects.

Chapter 7 investigates for the first time the availability of speaker indifference effects in the domain of modified numerals, both theoretically and experimentally. It proposes a tentative analysis of speaker indifference implications of modified numerals extending the composite theory from the previous chapter. It further reports a part of the second eye-tracking experiment that tests whether such an analysis would be on the right track. Although the processing of *at least* resembles that revealed in the case of speaker ignorance effects, the predicted processing patterns pointing to a strength difference in indifference effects between superlative and comparative modifiers were not attested. This difference between the two modifiers was not confirmed by off-line, judgement data either, thereby challenging further the suitability of the extended composite theory. Next, I ponder over the attested results, taking into account certain intuitions regarding speaker indifference effects of numeral modifiers, and discuss possible causes and directions for an explanation.

Finally, **chapter 8** brings together the results of the different chapters and reflects on how these contribute to a better understanding of the pragmatics of numeral modifiers as well as to the experimental pragmatics research in general.

CHAPTER 2

Theoretical background to numeral modifiers and their implications

2.1 Introduction

Geurts & Nouwen (2007) are the first to tackle the question as to whether the different truth-conditionally equivalent modified numeral expressions convey the same meaning, zooming in on superlative numeral modifiers (*at least, at most*) and comparative numeral modifiers (*more than, fewer/less than*). They argue that the key difference between these two types of modifiers is that the former unlike the latter have a(n epistemic) modal lexical meaning. This modal meaning is understood as a signal of uncertainty or ignorance on the part of the speaker, illustrated by the contrast that arises between (1-a) and (1-b) once we follow up with a sentence that reveals that the speaker has precise knowledge of the quantity in question.

(1) a. Magda ate more than three souvlakis. In fact, she ate five.
    b. Magda ate at least four souvlakis. # In fact, she ate five.

Soon a growing literature developed (with Büring, 2008 being the starting point) supporting that this epistemic component of superlative modifiers—or their *speaker insecurity* implication in Büring's words—is not part of their semantics, but it is rather derived via a pragmatic mechanism. Importantly, this has been the dominant take in the literature for almost a decade now, which also encompasses a number of hybrid proposals on the semantic/pragmatic encoding of speaker ignorance. Apart from pointing out the cross-linguistic

scope of the abundance of expressions referring to a particular quantity and of the distinction between superlative and comparative modifiers with respect to speaker ignorance inferences, Nouwen (2010) extends the distinction in question to bigger classes of modifiers, based on a considerable collection of armchair data. He proposes that numeral modifiers (across languages) are divided into two main classes, i.e., class A and class B: The latter class obligatorily gives rise to speaker ignorance (in non-embedding environments), while the former does not. In English, class A of numeral modifiers includes the comparative quantifiers *more/less/fewer than* as well as the locative prepositions *under/over* and *between ... and ...*, whereas class B includes the superlative quantifiers *at least/most*, *maximally/minimally*, disjunctions like *n or more/fewer* and the directional prepositions *from ... (up) to* and *up to*.

Although the study of superlative modifiers has monopolized the relevant literature, as theorists were seeking to elucidate the status of their ignorance implications, very recently some of the (pragmatic) accounts started turning their attention to comparative numeral modifiers (or class B modifiers more generally), too; especially since Westera & Brasoveanu's (2014) study, showing that comparative modifiers can also trigger ignorance effects if there is a precise question under discussion. We illustrate their point (supported experimentally) by the dialogue in (2). A's *how many* question in (2) asks for a precise answer, and the use of *more than* by B implies that B is unable to give the precise answer she is asked for, so she is ignorant about the exact number at issue.

(2) A: How many souvlakis did Magda eat?
B: She ate more than three souvlakis.
$\rightsquigarrow$ *B doesn't know exactly how many souvlakis Magda ate*

In section 2.2 that follows, I discuss four different views on numeral modifiers. We mostly focus on the derivation of speaker ignorance effects according to the different views, which for long have been thought to be the (only) defining difference between superlative numeral modifiers and comparative numeral modifiers. Section 2.2.1 is concerned with the semantic accounts of superlative modifiers and the speaker ignorance effects thereof, section 2.2.2 presents the pragmatic view whereby ignorance effects are derived via a Quantity-based reasoning, section 2.2.3 discusses more recent accounts that assume an inquisitive semantics for superlative numeral modifiers and derive their epistemic effects as a type of Quality implicature, while section 2.2.4 discusses accounts that treat numeral modifiers as modifiers of meta-speech acts. Section 2.2.5 summarizes the various accounts and brings attention to data concerning the semantic/pragmatic status and robust nature of speaker ignorance implications. Lastly, section 2.3 deals with the understudied variation effects of modified numerals and the proposals of some of the aforementioned accounts to capture those effects.

2.2 Different views on numeral modifiers: Speaker ignorance effects

2.2.1 The ignorance-as-entailment view

Geurts & Nouwen (2007) were the first to observe that, although the numeral phrases modified by the comparative and superlative quantifiers in examples (3) and (4) (repeated from above), respectively, have the same cardinal content[1], intuitively these examples communicate different information: While (3) conveys that the number of souvlakis Magda had is greater than 3, (4) seems to express that *the speaker is certain* that there is a group of four souvlakis and Magda ate each one of them, and *considers it is possible* that Magda ate even more (than four) souvlakis.[2]

(3) Magda ate more than three souvlakis.

(4) Magda ate at least four souvlakis.

Pointing out a number of (semantic) differences between comparative and superlative modifiers, Geurts & Nouwen conclude that the informative content of utterances like (3) and (4) differs in that the latter involves an epistemic modal component (note the italics above). They formalize their intuition as in (5) and (6) below (for (3) and (4), respectively), assuming that both comparative and superlative modifiers are focus sensitive (following Krifka, 1999) and as such they operate on the scale associated with their focused argument, i.e., the numeral in our case ((7) is the relevant scale for the examples above). (A note about the notations I will be using in the following—unless otherwise specified: I represent meanings by means of a simple language of first-order logic with pluralities and I use $\#x$ to indicate the cardinality of a plural individual x.)

(5) $\exists x[\#x > 3 \wedge souvlaki(x) \wedge eat(m, x)]$

(6) $\Box\exists x[\#x = 4 \wedge souvlaki(x) \wedge eat(m, x)] \wedge \Diamond\exists x[\#x > 4 \wedge souvlaki(x) \wedge eat(m, x)]$

(7) $< \ldots \lambda x[\#x = 5 \wedge souvlaki(x)],\ \lambda x[\#x = 4 \wedge souvlaki(x)],\ \lambda x[\#x = 3 \wedge souvlaki(x)] \ldots >$

In words, (5) says that there is a group of souvlakis with cardinality greater than three and Magda ate each of them, and (6), which faithfully reflects the authors' intuition too, expresses that the speaker is certain that there is a group of four souvlakis and Magda ate each of them, and further the speaker considers it possible that there is a group of souvlakis with a cardinality greater than four and Magda ate each of those. Although here we only exemplify the

[1] Assuming that only discrete numbers of souvlakis are under consideration in the current discourse.

[2] We are only concerned with the cases where focus is on the numeral.

lower-bound modifiers, the authors formulate a similar semantics for the corresponding upper-bound numeral modifiers. What is crucial in the present theoretical proposal is that the epistemic modal component that Geurts & Nouwen identify for superlative modifiers is placed in their lexical semantics. This implies the derivation of obligatory ignorance inferences, i.e., as entailments of unembedded uses of superlative modifiers.

Nouwen (2010), who gives a different semantics to comparative and superlative modifiers, also offers a semantic account of the derivation of speaker ignorance effects of sentences with superlatives, taking them to be the output of the compositional semantics of the relevant sentences. Let us see how he builds up his proposal. He argues that comparatives, and class A modifiers in general, express relations to a specific precise amount, while class B modifiers express relations to a range of values. Next, following Hackl (2000), he gives class A modifiers a degree semantics similar to that of other comparative constructions, and treats class B modifiers as maximum/minimum indicators.[3] Below, I illustrate the relevant semantics for our running examples with upper-bound modifiers.

(8) Magda ate fewer than five souvlakis.
$max_n(\exists x[\#x = n \wedge souvlaki(x) \wedge eat(m, x)]) < 5$

(9) Magda ate at most four souvlakis.
$max_n(\Diamond\exists x[\#x = n \wedge souvlaki(x) \wedge eat(m, x)]) = 4$

(8) says that the maximal value of having cardinality n such that Magda ate n-many souvlakis is smaller than 5, and (9) asserts that the maximal number of souvlakis Magda could have eaten, *according to the speaker's beliefs*, is four. Note the presence of the epistemic information in the interpretation of (9). According to Nouwen, the LF with the truth conditions in (9) without the (epistemic) $\Diamond$ operator is blocked by the corresponding LF containing a bare—rather than a modified—numeral given Grice's sub-maxim of Brevity, because the latter LF yields the same interpretation (*Magda ate four souvlakis*) but involves a much simpler form. Thus, the covert $\Diamond$ operator is inserted in order to rescue the derivation. Crucially, on the present account the import of this very operator is responsible for the speaker ignorance interpretation of the sentence in (9), which, as Nouwen notes, creates an epistemic range of values with four being the maximum. Although the import of this operator is pragmatically motivated, the resulting speaker ignorance is an obligatory inference derived via compositional semantics.

In the foregoing, I reviewed two semantic accounts of the speaker ignorance implication of modified numerals. Both accounts propose a different semantics for superlative and comparative numeral modifiers, but the one adds the

[3]In contrast to Geurts & Nouwen (2007), who formulate a semantics of an operator that associates with focus, Nouwen pursues a degree semantics analysis for comparative quantifiers similar to that in Hackl (2000), in order to make them parallel to other comparative constructions and, specifically, to capture their similar scopal behavior (cf. Heim, 2000).

epistemic component in the lexical meaning of superlative modifiers, while the other derives it via the application of a silent LF existential modal operator that licenses superlative modifiers in unembedded contexts. Also, in this kind of context the derivation of speaker ignorance inferences is obligatory under both accounts. Below, we are turning to pure pragmatic accounts of speaker ignorance.

2.2.2 The Quantity-based view

The most popular approach to speaker ignorance in the dominating pragmatic literature derives such implications as Quantity implicatures, with the majority of the relevant accounts drawing a parallel to similar implications of unembedded disjunction (cf. Gazdar, 1979; Sauerland, 2004, a.o.). Before presenting the details of the various pragmatic proposals, I would like to briefly introduce the standard (Quantity-based) recipe of implicatures (Geurts, 2011), which lies at the core of (the greater part of) those proposals.

Given an utterance ϕ that is a member of a (Horn) scale $< \psi, \phi >$, where ϕ and ψ are alternatives ordered in terms of entailment/informativeness, the first step of the standard recipe is the listener's observation that—given the aforementioned scale—the speaker could have made a stronger and more informative (and equally relevant) utterance ψ than her actual utterance ϕ. Then the listener reasons about the speaker's beliefs given her utterance choice as follows: The speaker believes what she said to be true (i.e., $\Box_{Bel}[\phi]$, schematically), but does not have the belief that any of the stronger (and relevant) alternatives are true, hence, $\neg\Box_{Bel}[\psi]$ (*primary Quantity implication*, Sauerland, 2004). This weak implicature is all we need for now. Take now again example (4), repeated below, and let us assume that the *at least* phrase is associated with a set of disjunctive alternatives coming from the (informativeness) scale in (10).

(4) Magda ate at least four souvlakis.

(10)

$$\left\langle \begin{array}{c} \text{exactly 4} \\ \text{more than 4} \end{array} , \text{ at least 4} \right\rangle$$

Applying the standard recipe described above to each of the alternatives, an utterance like (4) turns out to implicate that the (equally) stronger alternative propositions with *exactly* and *more than* in place of *at least* are not uttered by the speaker because she does not believe them to be true, that is, the speaker does not believe it is true that Magda ate exactly four souvlakis, i.e., $\neg\Box_{Bel}[4]$, and does not believe it is true that Magda ate more than four souvlakis, i.e., $\neg\Box_{Bel}[5, ...)$, (again, we take only discrete numbers of souvlakis into account). These implications together with the assertion lead to the conclusion that the speaker believes that the number of souvlakis Magda ate is in the range [4, ...), i.e., $\Box_{Bel}[4, ...)$, and lacks the belief that it is exactly 4 and that it is equal to any number greater than 4. Thus, she considers it possible that the number in

question is 4 and she considers it possible that it is 5 or more, hence, she is ignorant about the exact number of souvlakis.

This is the basic mechanism that Büring (2008); Cummins & Katsos (2010); Kennedy (2015); Nouwen (2015); Schwarz (2016a), and Spector (2015) share to some extent and implement in order to derive ignorance effects with superlative numeral modifiers. However, these accounts differ from each other in the design choices the authors make. More specifically, they differ in (i) the specification of the alternatives associated with superlative modifiers, (ii) the source of the alternatives, (iii) the kind and the nature of the belief operator $\Box_{Bel}$ assumed in the reasoning about the speaker's beliefs, (iv) the underlying assumptions they make about implicatures, and (v) the obligatoriness of drawing the implicature under discussion. In the following, I will elaborate more on (i)–(v) in relation to the relevant accounts.

(i) Specification of alternatives This concerns the specification of the alternatives assumed for the derivation of ignorance inferences with superlative modifiers. Accounts are split into those that posit disjunctive alternatives and those that include richer sets of alternatives. To start with Büring (2008), who was the first to make a parallel of superlative numeral modifiers with disjunction, he takes the meaning of *at least* to correspond to the disjunction *exactly n or more than n* at some level of description, and the relevant disjuncts to be fed into the Quantity-based recipe as the (relevant) stronger alternatives. Cummins & Katsos (2010), too, hold that superlative modifiers have a disjunctive meaning (at some level of representation), arguing that they are the natural language equivalents of the (non-strict comparison) operators $\geq$ and $\leq$, respectively, that bear a disjunctive status: i.e., $= n$ *or* $> n$ and $= n$ *or* $< n$.[4] Spector (2015) likewise takes superlative modifiers to be underlyingly (Hurford) disjunctions; *at least n* means "n or more" and is associated with the following set of alternatives: ALT = $\{$*at least n*, *exactly n*$\}$. Kennedy (2015) is the last to belong to the camp of disjunctive alternatives for utterances with *at least n*. On his account, the alternatives comprise the corresponding utterance with the bare numeral n, which he takes to have an *exact* semantics, and the corresponding utterance with a comparative modified numeral, i.e., *more than n*.

Let us now turn to accounts that assume that the scales associated with superlative modifiers consist in a more articulate set of alternatives as opposed to the stipulated set of disjunctive alternatives posited by the aforementioned pragmatic accounts. Schwarz (2016a) claims that the alternative utterances to an utterance with *at least n* that the speaker chooses not to make given

[4]It is noteworthy that, unlike Büring (2008), Cummins & Katsos explicitly adopt the classical generalized quantifier model (Barwise & Cooper, 1981) for numeral modifiers deriving the same truth conditions for comparatives and superlatives and, thus, without assuming any semantic difference between them. Yet, the fact that the latter quantifiers are further claimed to have a disjunctive status makes the authors talk about a so-called *augmented classical model*.

the standard recipe result from the interaction of the natural number scale (i.e., < ..., 3, 2, 1 >) and a scale that *at least* forms with *only*. The latter is determined based on the syntactic substitutability (Horn, 1972) of *at least* by other lexical items and its distribution (see also Mendia, 2016a for a very closely related proposal for *at least* as a numeral modifier).[5] To illustrate, the stronger alternative propositions for (4) are the following: [4], [5], ... (where again $[n]$ stands for the proposition with *only* n) and $[5, ...)$, $[6, ...)$, ... (where $[n, ...)$ stands for the proposition with *at least* n). Nouwen (2015) assumes the exact same resulting sets of alternatives as in Schwarz (2016a) for *at least*, but analogizes class B numeral modifiers to epistemic indefinites as opposed to disjunction, which as we will see next gives rise to non-stipulative alternatives (to some extent) in contrast to the alternatives in all previous accounts. But where do all these alternatives come from? This is what (ii) refers to.

(ii) Source of alternatives As already mentioned, Büring (2008), Cummins & Katsos (2010), and Spector (2015) take the meaning of superlative numeral modifiers to be disjunctive at some level of representation, with the latter author being more explicit about that giving *at least* the semantics of a disjunction. In all three accounts, then, the relevant disjuncts constitute the (stronger) alternatives of superlative modified numerals. Kennedy (2015), on the other hand, takes his disjunctive alternatives to utterances with superlative modified numerals to come from Horn scales (Horn, 1972). More accurately, he assumes that utterances with numerals are associated with Horn scales that contain all relevant utterances with bare/unmodified and modified numerals as alternatives. The following ordering relations hold as far as superlative modifiers are concerned: Utterances with superlative modified numerals are asymmetrically entailed by the alternative utterances with bare numerals and by the alternatives with the corresponding comparative modified numerals. Given that comparative modifiers are part of such scales (see (10)) with these particular entailment/ordering relations, no ignorance effects are expected to arise with those modifiers. This is so because neither the bare/unmodified (or *exactly*) alternative nor the superlative alternative entails the relevant utterance with the comparative, which are necessary according to the standard recipe. Schwarz's (2016a) two scales of alternatives, too, are Horn scales where alternatives are ordered by entailment/informativeness.

Finally, Nouwen's (2015) alternatives to an utterance with a superlative modified numeral like *at least* n (as in (4)) come from two different sources: The *exactly* (or *only*) alternatives, namely, the propositions [4], [5], ..., come from the anti-specificity presupposition he posits for superlative modifiers and the remaining alternatives, that is, the propositions $[5, ...)$, $[6, ...)$, ..., stem from

[5]Schwarz (2016a) draws heavily on Mayr (2013), who also considers two sets of alternatives, with the difference being that Mayr has *at most* in place of *only* in the second scale. However, Mayr's (2013) account focuses on the scalar implicatures triggered by modified numerals in embedded contexts and fails to deal with the ignorance implications *at least* seems to be associated with in unembedded contexts.

substituting (only) the numeral with stronger numerals from the Horn number scale, as in Schwarz (2016a). Let us see what Nouwen's anti-specificity presupposition is and what reasoning it can lead to. Nouwen (2015) makes a parallel of superlative modifiers—and class B modifiers in general—with epistemic indefinites (see Alonso-Ovalle & Menéndez-Benito, 2010). Similarly to the Spanish epistemic indefinite *algun*, superlative modifiers carry an anti-specificity presupposition (or lexical requirement) such that the domain of the cardinality of the set they express is not a singleton, i.e., there is no specific number or single option in that domain. Given an utterance with a superlative modifier like (4) signaling that the domain of quantification is not a singleton, the listener might reason as follows: Why did the speaker use *at least*, an expression that signals that the cardinality at issue cannot be specified or expressed by a single option/exact number, rather than using a stronger alternative with a singleton domain, such as an *exact* expression, i.e., *exactly four/exactly five/exactly six/... souvlakis*? The listener concludes that the speaker made this choice because no *exactly* alternative is true. Hence, the following implicatures arise: $\neg[4]$, $\neg[5]$, $\neg[6]$, ... Applying the standard informativity-based reasoning (sketched at the beginning of this section) to the rest of the alternatives (i.e., Horn-scale alternatives) and combining the resulting implicatures with the aforesaid presupposition-based ones and with the assertion $\Box_{Bel}[4, \ldots)$ derives the speaker ignorance implications: The speaker considers it possible that [4] and she considers it possible that $[5, \ldots)$. Crucially, Nouwen (2015), contrary to all aforementioned theorists, motivates the *exactly* alternatives he assumes for superlative modified numerals via the anti-specificity requirement.

(iii) The belief operator As regards the stage of the listener's reasoning in the standard recipe that has to do with the speaker's beliefs about the relevant alternatives, this is summarized by the operator $\Box_{Bel}$, the main issue addressed in this paragraph. Specifically, $\Box_{Bel}$ has various disguises across accounts: It can have the form of a personal/speaker-oriented epistemic certainty operator (Büring, 2008; Cummins & Katsos, 2010; Nouwen, 2015; Schwarz, 2016a; Spector, 2015) or of an impersonal epistemic certainty operator (Kennedy, 2015). In addition, it can be an LF operator as in Nouwen (2015), and Spector (2015), being a necessary part of the assertion, or it can be part of the pragmatic reasoning (Cummins & Katsos, 2010; Kennedy, 2015), sometimes explicitly supplied through the Quality maxim (Büring, 2008; Schwarz, 2016a). These differences do not matter for our current purposes.

(iv) View on implicatures The aspect highlighted here pertains to how ignorance implicatures are taken to be generated and to the general understanding of what an implicature is. A basic distinction then applies to the relevant accounts, specifically exhibited between those accounts that adopt a grammatical approach to Quantity implicatures (Chierchia, 2006; Chierchia et al., 2013) and those that fall into the (neo-)Gricean pragmatic tradition

(e.g., as in Geurts, 2011; Sauerland, 2004). While the latter apply the standard recipe as presented at the beginning of the present section as far as primary implicatures are concerned, the former deviate from that in that the reasoning about the relevant (stronger) alternatives and the speaker's beliefs with respect to those happens in the grammar. In other words, the computation of Quantity implicatures on this view is a regular compositional or grammatical process. The main operator of this process is the so-called exhaustification operator, which is the silent equivalent of *only*, as it exhaustifies an assertion, by negating the relevant alternatives. To illustrate, say we take sentence (4), repeated below, and a disjunctive set of alternatives; then the resulting interpretation after the application of the grammatical exhaustification operator O is schematically given in (11), notice the similarity to the standard Gricean recipe.

(4) Magda ate at least four souvlakis.

(11) $O_{ALT}(\Box_{bel}[4, ...))$
$= \Box_{Bel}[4, ...) \wedge \neg\Box_{Bel}[4] \wedge \neg\Box_{Bel}[5, ...)$

That is, this ignorance interpretation is the same as in the main mechanism I presented in the very beginning, i.e., the speaker believes that the number in question is in the range [4, ...), but does not believe that it is equal to 4 and does not believe that it is 5 or more. This means that one arrives at the interpretation we are after, no matter what view of Quantity implicatures they adopt. Among all Quantity-based accounts discussed in this section only Spector (2015) explicitly formulates his account within a grammatical framework.

(v) Obligatoriness of ignorance The last aspect that differentiates the relevant accounts is to what extent it is obligatory for the listener to draw the ignorance implicature when the speaker makes an utterance with a superlative modified numeral like that in (4) (this aspect will be further discussed and will play a central role in the experiments reported in chapters 5 and 6). Spector (2015) specifies for unembedded occurrences of *at least* that the relevant speaker ignorance implicature is derived via the application of an obligatory exhaustification operation above the speaker belief operator he assumes propositions are prefixed with (i.e., Meyer's (2013) K operator standing for *the speaker believes that*).[6] As a result, when *at least* appears in non-embedding declarative environments, it necessarily gives rise to speaker ignorance. Büring (2008), Cummins & Katsos (2010), Kennedy (2015), and Schwarz (2016a) do not make such a specification for the application of the standard Quantity-based recipe they adopt, suggesting an optional character for the speaker ignorance implicature they derive.

The aspect in question is further left unclear in Nouwen's (2015) account.

[6] As he explains *[a]t least is vacuous [...] unless the alternatives it introduces are at some point used. This is what forces exh (Brevity constraint)* (Spector, 2015, p. 14), where *exh* stands for the exhaustification operation.

On this account, superlative modified numerals (and class B modified numerals in general) carry an anti-specificity presupposition such that the domain of the cardinality of the set they express is not a singleton. Given the specifics of this account, it is predicted that an *at least*-utterance can felicitously be performed by a knowledgeable speaker. For instance, on this account, (4) asserts that the cardinality of (the set of) souvlakis Magda ate is in the range $[4, ...)$ and presupposes that the domain of this cardinality has more than 1 option/number. That is, it presupposes that the domain does not contain a single option/number like, e.g., $[5]$, but multiple options, as in the case of the subset $[4, 5, 6]$, for example. So in the case of the latter set, a speaker uttering (4) is taken to believe that the cardinality of souvlakis Magda ate is in the set $[4, 5, 6]$, which is compatible with the speaker actually being knowledgeable and, say, (also) believing that the cardinality in question is $[5]$, as this is just a subset of $[4, 5, 6]$. Therefore, *at least* can felicitously be uttered by a knowledgeable speaker on Nouwen's account. However, the use of a numeral with an anti-specificity presupposition *raises the issue why a singleton domain is excluded by the speaker* in Nouwen's words (Nouwen, 2015, p. 261–262). This initiates a reasoning about the (stronger) alternatives of the scale defined based on the specificity/width of the relevant domains, leading to speaker ignorance (remember that another type of speaker ignorance via reasoning about alternatives based on the number scale is calculated on top). Importantly, it is not clear whether this issue is obligatorily raised once an expression with an anti-specificity lexical requirement is used, which would result in obligatory speaker ignorance. But why would a speaker pick an expression like *at least* if its presupposition is not to be used in some way, e.g., triggering an implicature? Although it remains an open issue whether Nouwen (2015) wants the implicature mechanism to apply obligatorily, his presupposition-based proposal seems to perfectly capture the robustness of speaker ignorance implicatures with superlative modifiers as opposed to comparative modifiers, and their cancellability resistance, to be discussed in section 2.2.5. Note that Nouwen (2015) does comment that the ignorance implicatures triggered by alternatives based on the Horn number scale might be more defeasible than those triggered based on his anti-specificity presupposition (e.g., *exactly* alternatives), implying a greater strength for the latter.

The accounts included in the present section cover the same facts as far as speaker ignorance implications are concerned, regardless of the individual design choices reported in (i)-(iv). By that I mean to say that (eventually) they derive the same ignorance implication for an *at least* n-utterance such that $[n]$ and $[n + 1, \ldots)$ are both epistemic possibilities. Perhaps the aspect discussed in (v) concerning the obligatoriness of drawing an ignorance implicature makes the proposals under consideration differ to some extent in empirical coverage. As was already suggested in the introduction of this chapter (see contrast in (1) in section 2.1) as well as in the previous paragraph, which I will further allude to in section 2.2.5, speaker ignorance inferences seem not to be defeasible. Re-

latedly, some of the proposals explicitly account for the robustness of ignorance implicatures by motivating the obligatoriness of the pragmatic mechanism they adopt (e.g., Spector, 2015), while others leave this point implicit (cf. Nouwen, 2015) or even untouched (pure neo-Gricean accounts). It is certain, though, that the various accounts exhibit differences in empirical coverage due to the specific design choices they make with regard to other aspects of numeral modifiers, e.g., degradability of superlative modifiers with negation. Two of the five aspects we discussed in this section will be relevant and crucial for the present thesis, also in relation to the remaining theoretical accounts. More specifically, the source of alternatives and the obligatoriness, as already noted, of speaker ignorance inferences—two aspects that are closely related—will play a key role in our ignorance experiments to be presented in chapters 5 and 6.

2.2.3 The Quality-based view

After considering the Quantity-based approach to the ignorance effects of superlative numeral modifiers, we will next be concerned with a couple of "hybrid" accounts of speaker ignorance, in the sense that those accounts involve both a semantic/lexical, or more conventional, and a pragmatic encoding of speaker ignorance. Those accounts offer an analysis of superlative as well as of comparative numeral modifiers, which will be both presented below with a special focus on the derivation of speaker ignorance effects.

First, we will go through Coppock & Brochhagen's (2013b) account of superlative and comparative modifiers, who formulate their account in inquisitive semantics. On this account, utterances with superlative modifiers, unlike utterances with comparative modifiers, have an inquisitive content on top of their informative content, which is equivalent to that of comparative modifiers. That is, an utterance with *at least*, for instance, is taken to introduce multiple semantic alternatives (alternative propositions), which means that it expresses an issue as to which of these alternatives holds, hence the term *inquisitive*. An utterance with *more than*, on the other hand, generates a set containing a single semantic alternative, thus, it has no inquisitive content. More specifically, the set of alternatives introduced by a sentence with *at least* contains the alternative propositions that are ranked as high or higher on a pragmatically defined scale: i.e., for (4) this set is $\{[4,...),[5,...),[6,...),...\}$ assuming a one-sided reading of numerals and $\{[4],[5],[6],...\}$ assuming a two-sided analysis of numerals (note that $[n]$ stands for the proposition with an *at least* reading of n, and $[n,...)$ for the proposition with an *exact* reading of n). Besides the standard Gricean Quality maxim (*Make your contribution one that is true*), Coppock & Brochhagen posit an additional Quality maxim, the so-called *Maxim of Interactive Sincerity*, which instructs the following: *Don't make an inquisitive utterance if you already know how to resolve the issue your utterance expresses.* In other words, if a speaker makes an (multiple) alternative-generating utterance, her information state should support such an utterance. An utterance with a superlative modifier (irrespective of the type of analysis we assume for bare numerals) in

combination with this Quality maxim gives rise to an ignorance implicature. For (4) that would be that the speaker is ignorant about the exact number of souvlakis in the range $[4, ...)$ Magda ate. No such implicature arises with comparative modifiers, as the relevant utterances do not generate multiple alternative propositions, hence they are not subject to the Quality maxim of Interactive Sincerity.

Notice, first, that the ignorance derived on the basis of the semantics of superlative modifiers on this account in combination with the particular Quality maxim assumed on top gives rise to an obligatory type of speaker ignorance. To elaborate, the content of this Quality maxim is very closely tied to the inquisitive semantics of superlative modifiers, which makes it hard to violate such a maxim. Also, Quality implicatures are known to be robustly derived compared to other types of implicatures, such as Quantity or Manner implicatures. As Grice puts it in the *William James Lectures*:

> It is obvious that the observance of some of these maxims is a matter of less urgency than is the observance of others; a man who has expressed himself with undue prolixity would, in general, be open to milder comment than would a man who has said something he believes to be false. (Grice, 1989, p. 27)

There is also strong experimental evidence indicating the robustness of Quality implicatures. In a series of experiments, Pantazi, Kissine & Klein (in press) find that their participants believe messages they receive even though they are given explicit meta-information signaling the messages' falsity.[7] In line with this, an utterance with *at least* is highly unlikely to be performed by a speaker that is knowledgeable of the cardinality under discussion, as it is hard to cancel the Quality-based ignorance implicature it is associated with. On the contrary, this is not the case for the corresponding utterance with a comparative modifier, as this is not subject to the Quality maxim in question.

The next thing to point out is that the speaker ignorance inference derived by Coppock & Brochhagen's (2013b) account is different from the one derived via the Quantity-based accounts presented in the previous section. To illustrate with sentence (4), the ignorance meaning the Quantity-based accounts derive says that the speaker believes that the number of souvlakis Magda ate is in the range [4, ...) (standard Quality implication) and the speaker does not believe that this number is 4 and she does not believe that it is greater than 4 (primary Quantity implications). On the other hand, Coppock & Brochhagen's (2013b) account derives the following ignorance meaning: The Maxim of Interactive Sincerity implies that the speaker's information state cannot support an utterance that expresses an exact quantity, which combined with the standard Quality assumption $\Box_{Bel}[n, ...)$ gives for (4) that the speaker believes that the number of souvlakis Magda ate is in the range $[4, ...)$ and she lacks any beliefs about (exact) numbers in that range. This is actually the main critical point Schwarz (2016b) raises with respect to Coppock & Brochhagen's (2013b)

[7] Thanks to Myrto Pantazi for discussion on that.

inquisitive semantics account. Alexandropoulou, Dotlačil, McNabb & Nouwen (2015) raise a similar point regarding the variation/free-choice-like effects of *at least* given Coppock & Brochhagen's (2013b) account (cf. also more extensive consideration in section 2.3.1 and chapter 3). Relatedly, Mendia (2016c) argues and shows experimentally that the ignorance inferences of superlative numeral modifiers *must* include the exhaustive interpretation of the prejacent as an epistemic possibility for the speaker. Namely, the inference that the speaker considers it possible that Magda ate *exactly/only four* souvlakis for (4). (In chapter 3, I will show that the corresponding inference is a necessary part of the variation implicature *at least* triggers in embedding environments.)

From now on, I will distinguish two types of speaker ignorance:

(12) a. *underspecified ignorance*: the speaker believes that the quantity is in $[n, \ldots)$, but doesn't hold any beliefs about numbers in that range
b. *specific ignorance*: the speaker considers it possible that $[n]$ and she considers it possible that $[n+1, \ldots)$.

Underspecified ignorance is what Coppock & Brochhagen (2013b) derive, while the Quantity-based accounts derive specific ignorance. This distinction will play an important role in the ignorance experiments in chapters 5 and 6. Note also that the speaker ignorance derived based on the alternatives only coming from the anti-specificity presupposition in Nouwen's (2015) account is a type of what we just termed underspecified speaker ignorance. As will become clear in section 2.3.1, which type of ignorance we derive depends on the specification of alternatives we make. This is yet another aspect that will be relevant in the present thesis, specifically in the experimental investigation of variation inferences in chapter 3.

Ciardelli, Coppock & Roelofsen (2017) put forth an amendment of Coppock & Brochhagen's (2013b) account, also in inquisitive semantics. Besides a framework issue they settle, their account improves on two main aspects: (i) They aim to capture the specific speaker ignorance inference as opposed to the underspecified one discussed above, and (ii) they aim to accommodate Westera & Brasoveanu's (2014) recent experimental finding that comparative numeral modifiers, too, can trigger speaker ignorance effects. As in Coppock & Brochhagen (2013b), utterances with *at least* generate multiple alternatives, but unlike Coppock & Brochhagen (2013b), yet similarly to Büring's (2008) original idea, such utterances have a disjunctive semantic meaning, illustrated here for (4): {[4], [5, ...)}. As before, *more than* introduces a set of a single semantic alternative, i.e., {[4, ...)} for the equivalent comparative version of (4) given below repeating (3).

(3) Magda ate more than three souvlakis.

Furthermore, *at least* and *more than* are taken to have similar lexically determined pragmatic alternatives, similar to those proposed by Schwarz (2016a)

for *at least*, which are available for Quantity-based reasoning:

(13) *at least n*: $\{\textit{at least } m \mid m \in \mathbb{N}^*\} \cup \{m \mid m \in \mathbb{N}^*\}$
more than n: $\{\textit{more than } m \mid m \in \mathbb{N}^*\} \cup \{m \mid m \in \mathbb{N}^*\}$

It is essential here to note that Ciardelli et al. (2017) establish a distinction between lexically determined pragmatic alternatives and contextual pragmatic alternatives: The latter are the lexical pragmatic alternatives that are relevant to the question under discussion (QUD). Like Coppock & Brochhagen (2013b), they posit an additional Quality maxim, the *Inquisitive Sincerity* maxim, which states a similar requirement to Interactive Sincerity: *if ϕ is inquisitive, then the speaker's information state s should not already resolve the issue expressed by ϕ: $s \notin [[\phi]]$* (Ciardelli et al., 2017, p. 10). So for the inquisitive utterance (4), repeated below, the Gricean Quality maxim requires that the speaker believes $\{[4], [5, ...)\}$ to be true and the Inquisitive Sincerity maxim, as in Coppock & Brochhagen (2013b), requires that the speaker's information state $s \nsubseteq [4]$ and that $s \nsubseteq [5, ...)$, namely, the speaker should not believe that the number of souvlakis Magda ate is exactly four, nor should she believe that the number of souvlakis Magda ate is greater than four.

(4) Magda ate at least four souvlakis.

Combining the standard Quality implication with the above one, we arrive at the specific ignorance implicature that the speaker considers it possible that the number of souvlakis Magda ate is exactly 4 and that she considers it possible that this number is greater than 4. Quality-based ignorance implications are always available for *at least*, regardless of the question under discussion, as they are derived based on the semantic alternatives it generates. Such implicatures are not available with *more than* utterances, because they generate a single semantic alternative and thus are not subject to the Inquisitive Sincerity maxim.

Let us see, however, what the present account predicts if we move on to a standard Quantity-based reasoning. Especially for the case that (3) is uttered in a context with a precise (*how many*) question under discussion (see (2) repeated below), which has been shown experimentally to facilitate ignorance inferences with *more than* (Westera & Brasoveanu, 2014).

(2) A: How many souvlakis did Magda eat?
B: She ate more than three souvlakis.

First, we should consider which of the lexical pragmatic alternatives given in (2) are relevant to a *how many* question, namely, which ones qualify as contextual pragmatic alternatives. The *how many* question in (2) is associated with this set of alternatives: $\{[0], [1], [2], [3], [4], \ldots\}$. Hence, the lexical pragmatic alternatives $\{\textit{more than } n \mid n \geq 4\} \cup \{(\textit{exactly})\ n \mid n \geq 4\}$ of B's utterance in (2) all qualify as relevant alternatives to the question under discussion, and none of them is inquisitive. Thus, so far we have derived just the Quality impli-

cation that B sincerely uttered {[4, ...)}, that is, B believes that the cardinality of souvlakis Magda ate is greater than three. Given that B did not sincerely utter any of the above (non-inquisitive) pragmatic alternatives, which would have been more informative, we infer that B doesn't know whether Magda ate exactly four souvlakis and that B doesn't know whether the number of the souvlakis Magda ate is greater than 4. This combined with what the speaker has sincerely uttered signals that B considers it possible that Magda ate exactly four souvlakis and she considers it possible that Magda ate more than four souvlakis. Notice that this is exactly the speaker ignorance implicature that the Quantity-based accounts derive for *at least*-utterances, with the difference being that here it is tied to the question under discussion. Since the lexical as well as the contextual pragmatic alternatives for *at least* are the same as for *more than*, the same inference is drawn for *at least* at the Quantity-related phase of pragmatic reasoning, which is identical to the inference we derived above given the Quality maxims. Therefore, nothing new is implicated for *at least* and nothing different from what is implicated for *more than*.

In spite of the resemblance of the inference derived for *at least* and *more than* utterances, the authors do make a distinction as regards the strength of the inference in question: They hold that the speaker ignorance inference of the former utterances are more robust compared to the latter, because (i) they are derived via a Quality maxim, whose observance is of greater importance and urgency compared to a Quantity maxim or any other maxim, as has already been highlighted by Grice (1989), and (ii) they are also derived via Quantity-based reasoning on the relevant lexical pragmatic alternatives. This account implicates that speaker ignorance implicatures are available with *at least* utterances in various contexts and are obligatory, because of (i) and (ii), and because they are tied to the Quality maxim of Interactive Sincerity by virtue of the inquisitive semantics of *at least* (cf. similar discussion for Coppock & Brochhagen, 2013b). Given that, *at least* is unlikely to be uttered by a knowledgeable speaker whose information state resolves the issue expressed by the relevant utterance (cancellation-resistance of Quality implicatures contrary to other types of implicatures).

In this section, we saw accounts that derive the implication that the speaker does not believe $[n]$ is true and she does not believe $[n+1,\ldots)$ is true for an *at least* n utterance via a Quality-based reasoning. This results in a speaker ignorance implication of a more obligatory nature compared to a Quantity-based ignorance implication like that derived by the accounts we considered in the previous section (note that given Nouwen, 2015 the first part of the aforesaid implication might be considered to be more robustly available than the second part due to the different source of the corresponding alternatives).

2.2.4 Speech act views

There is a number of proposals that cannot be classified with respect to the standard truth-conditional semantics-pragmatics divide, generally adopted by

the accounts discussed in the foregoing. These proposals I am referring to make use of the notion of speech acts and in fact bring in a quite different perspective, whereby what we have called so far numeral modifiers are taken to operate at a higher level, that of speech acts, and, more precisely, they are treated as modifiers of (basic or meta-) speech acts. Although I chose to discuss those accounts in a separate section by virtue of their framework commonality, as the reader will notice in the following, they clearly exhibit some similarities to the accounts presented in the previous sections, in particular as concerns the derivation of speaker ignorance implications of superlative modified numerals.

We will start with the account by Cohen & Krifka (2014) (see also previous work in Cohen & Krifka, 2010). Cohen and Krifka take superlative modifiers to express meta-speech acts, and more specifically the meta-speech act GRANT. If a speaker GRANTs a proposition ϕ, she refuses to assert the negation of ϕ, i.e., $\neg$ASSERT($\neg\phi$). So a sentence with *at least*, like (4), expresses that the minimal number n, such that the speaker GRANTs that Magda ate n souvlakis, is 4. This is further interpreted as saying that for any $n < 4$ the speaker does not GRANT that Magda ate n-many souvlakis, that is, the speaker excludes the following assertions:

(14) Magda ate exactly three souvlakis.
Magda ate exactly two souvlakis.
Magda ate exactly one souvlaki.
Magda ate exactly zero souvlakis.

Cohen and Krifka note that excluding assertions, as we did above, is equivalent to making the following ones, respectively:

(15) Magda did not eat exactly three souvlakis.
Magda did not eat exactly two souvlakis.
Magda did not eat exactly one souvlaki.
Magda did not eat exactly zero souvlakis.

Hence, an utterance with a superlative modifier expresses which propositions the speaker asserts to be false, but not which ones she considers to be true. That is, by such an utterance the speaker states the falsity-conditions of a proposition, while the truth-conditions remain unknown/undefined. So, for example, we would not know what the truth of (4) is given a situation where Magda ate exactly five souvlakis, since the basic interpretation the present account gives to (4) leaves this open, as it just excludes the assertions in (14) and is compatible with the speaker also asserting that Magda did not eat exactly five souvlakis. So we need to exclude assertions like *Magda did not eat exactly four souvlakis* or *Magda did not eat exactly five souvlakis*, etc., in order to determine the truth-conditions of (4). Cohen and Krifka suggest that we obtain the truth-conditions of (4) via Quantity-based reasoning. The speaker

uttering (4) has committed to all assertions in (15), where the number n of souvlakis Magda ate is smaller than four. If the speaker wanted to or could commit to the assertion that Magda did not eat five souvlakis, she should have done so according to the standard Quantity maxim. The fact that she didn't do so signals that the speaker was not committed to such an assertion or to any other similar assertion where $n > 5$. Hence, for any $n \geq 4$ the speaker considers it possibly true that Magda ate n-many souvlakis, which can be interpreted as indicating ignorance on the part of the speaker with respect to exactly how many souvlakis Magda ate.

Note that the speaker ignorance interpretation taken to arise with *at least* under the present account is different from the underspecified speaker ignorance derived by Coppock & Brochhagen (2013b) or the specific speaker ignorance derived by Ciardelli et al.'s (2017) Quality-based account and by Quantity-based proposals. Cohen and Krifka predict that the speaker considers possible *any* number in the relevant range, namely, they derive total speaker ignorance (see, e.g., Alonso-Ovalle & Menéndez-Benito, 2013 on total ignorance implications). As they further point out in relation to their example in (16), *this implicature may get weaker the higher the numbers get, because there might be additional reasons why the speaker might consider it impossible that John petted, say, 1,000 rabbits* (Cohen & Krifka, 2014, p. 60).

(16) John petted at least three rabbits.

Lastly, deriving the relevant ignorance implicature appears to be obligatory in the present account, as it is necessary for accounting for the truth of the *at least*-sentence. Therefore, a knowledgeable speaker is not expected to utter such a sentence.

Spychalska's (2015) account too takes both superlatives and comparatives to be modifiers of speech acts, of assertions in particular. The basic trait of her proposal is that she distinguishes between the truth-conditions and the assertibility conditions of an expression. She takes comparative and superlative modifiers to have the same truth-conditions, but different assertibility conditions. She draws a parallel between assertibility conditions of superlatives and disjunctions: Superlative sentences have modal assertibility conditions, which are also responsible for their epistemic effects. On the contrary, the assertibility conditions of comparative expressions have no such modal component. Assertibility conditions divide into the belief condition and the closure condition. The belief condition of an utterance with *at least* n requires that the speaker considers two possibilities: The possibility of n and the possibility of *more than* n (notice again the disjunctive note as well as the similarity to Ciardelli et al.'s (2017) proposal). So the belief condition of (4), repeated below, requires that the speaker considers it possible that Magda ate exactly four souvlakis and she considers it possible that Magda ate more than four souvlakis, which is the speaker ignorance implication that most of the Quantity-based accounts in the previous section derive.

(4) Magda ate at least four souvlakis.

The closure condition requires that the speaker does not consider it possible that $n < 4$ where n the number of souvlakis Magda ate in (4). On the other hand, the belief condition for *more than* utterances just requires that the speaker considers it possible that the number n in question is greater than m ($n > 3$ for the comparative equivalent of (4)) and the closure condition requires that the speaker does not consider it possible that $n \leq 3$ in the relevant example, hence, the speaker (only) considers that it is true that $n > 3$.

Given the above, speaker ignorance only arises with superlative modifiers as opposed to comparatives, and it actually arises obligatorily, as it is connected with the modal assertibility conditions of superlative modifiers. In other words, if a knowledgeable speaker informed about the exact number of souvlakis Magda ate utters (4), infelicity will arise due to violation of the belief condition in particular. Notice the parallel of assertibility conditions with a Quality maxim, as they both pertain to speaker's beliefs. Given the robustness of Quality maxims discussed in section 2.2.3, I take assertibility conditions to be equally hard to be violated. The inescapability and robustness of assertibility conditions can also be inferred by Spychalska's (2015) claim that they can even influence the production and evaluation of truth-conditionally valid inferences.

In this section, we looked at speech act accounts of modified numerals and of their speaker ignorance inferences. The ignorance implications these accounts derive are of an obligatory nature, similarly to the Quality-based accounts and unlike the Quantity-based accounts described in the previous sections.

2.2.5 Discussion

To recap, section 2.2 reviewed various theoretical proposals on numeral modifiers focusing on superlative and comparative modifiers, and on (what was initially determined as) the key difference thereof, that is, the speaker ignorance signal conveyed by the former. We saw accounts that encode this speaker ignorance signal semantically, either as part of the lexical semantics of superlative modifiers (Geurts & Nouwen, 2007) or as the (by)product of the compositional semantics of a sentence with a superlative modifier (Nouwen, 2010), and in both cases assuming a different semantics for comparative and superlative modifiers. Next, we turned to pragmatic accounts that derive ignorance implications via standard Quantity-based reasoning, most of which draw a parallel between superlative modifiers and disjunction (Büring, 2008; Cummins & Katsos, 2010; Kennedy, 2015; Spector, 2015) or epistemic indefinites (Nouwen, 2015), and assume either a binary set of alternatives (Büring, 2008; Cummins & Katsos, 2010; Kennedy, 2015; Spector, 2015) or more articulate sets of alternatives (Nouwen, 2015; Schwarz, 2016a), whereby alternatives are introduced at the LF or pragmatically via lexically determined Horn-scales or a disjunctive interpretation. I further reported some hybrid accounts of speaker

ignorance adopting the inquisitive semantics framework (Ciardelli et al., 2017; Coppock & Brochhagen, 2013b), where superlative modifiers are taken to have an alternative-introducing semantics and ignorance implications arise as Quality implicatures, via reasoning given certain types of Quality maxims on top of the Gricean one. One of those also included an account of speaker ignorance effects with comparative modifiers (Ciardelli et al., 2017), which are argued to trigger such effects when they are uttered as answers to a question that asks for a precise answer. We finally saw accounts that take numeral modifiers to actually be modifiers of speech acts, either deriving speaker ignorance effects of superlatives via Quantity-based reasoning or by means of some sort of condition on their use, similar to the Quality maxims/conditions assumed by the aforementioned inquisitive semantics accounts.

As has become clear so far, the dominant view on the derivation of speaker ignorance effects has it that they come about via some sort of pragmatic reasoning, either exploiting a Quantity maxim or a Quality maxim. Indeed, taking into consideration pragmatic diagnostics like embedding in downward entailing environments (see (17)) or reinforceability (see (18)), we can conclude that speaker ignorance has the status of an implicature, as it seems to disappear (or weaken) in (17), whose (dominant) reading is that *in all worlds where the number of people at the party was* ≥ 50*, the party was a success*, and it is reinforced by the *but* continuation sentence in (18) without redundancy.

(17) If at least 50 people came to the party, the party was a success.

(18) At least 50 people came to the party, but I don't really know how many exactly.

Although speaker ignorance did not fail the previous tests, it is perhaps hard to maintain the same as straightforwardly as far as the cancellability test is concerned; see (19), where the continuation sentence is supposed to cancel the ignorance signal of the sentence with *at least*.

(19) At least 50 people came to the party. ??Actually, to be precise, there were 53 people at the party.

(20) More than 50 people came to the party. Actually, to be precise, there were 53 people at the party.

As also suggested in the introduction, although the continuation sentence in (19) is not completely illicit, ignorance inferences seem to do quite bad at what is often thought to be the most common diagnostic of conversational implicatures (cf. contrast with (20)). In the following, I will focus on this very observation.

The apparent strength or persistence of ignorance implications is perhaps one of the reasons why the existing literature is so divergent, as we see almost anything from genuine semantic approaches to pure pragmatic approaches to speaker ignorance. We should note that, apart from the semantic accounts,

accounts that include a somewhat conventional basis for speaker ignorance inferences combined with a pragmatic reasoning, such as Ciardelli et al. (2017); Cohen & Krifka (2010); Coppock & Brochhagen (2013b); Nouwen (2015); Spector (2015); Spychalska (2015), can very well capture (19). Recall that we have highlighted that in these accounts speaker ignorance implications are predicted to emerge robustly with such utterances. Having said that, one could argue that it is not surprising that ignorance effects of superlative modifiers exhibit such a controversial status. If we consider the literature of epistemic indefinites (or differently, existential free choice indefinites in Chierchia's (2013) terminology) or of unembedded disjunction, there, too, ignorance effects are shown to be rather robust implications. The infelicity of the continuation sentences in (21) and (22) below illustrate this point, using the symbols # or ?? at the front.

(21) María se casó con algún estudiante del departamento de
María SE married with ALGÚN student of the department of
lingüística: # en concreto con Pedro.
linguistics: namely with Pedro
'María married a linguistics student, namely Pedro.'
(adapted from Alonso-Ovalle & Menéndez-Benito, 2010)

(22) John is in London or he is in Paris. ??In fact, he is in Paris.
(adapted from Lauer, 2013, p. 263)

These observations could further be taken to suggest that accounts that analyze superlative modifiers on a par with disjunction or epistemic indefinites, using a similar underlying implicature-mechanism, are on the right track. That is, the infelicity arising in (16) is the same as in (21) and (22), so whatever accounts for the former accounts for the latter too.

Shedding light on the dubious (semantic/pragmatic) status of ignorance effects of superlative modifiers must have been the main drive of the research on modified numerals to turn sooner or later to the collection and study of experimental data. The use of experimental methods rather than plain diagnostics like those applied above can help determine whether ignorance effects are semantic inferences or they are derived via some pragmatic process, which by now seems to be a puzzle. That is, well-controlled experiments with designs that have effectively been used in the field of experimental semantics/pragmatics can more systematically and thoroughly test for speaker ignorance availability and its status compared to informal methods collecting individual data points (see Sprouse, Schütze & Almeida, 2013 and references therein for such a comparison between methods). Establishing the status of speaker ignorance inferences is crucial for concluding what a suitable analysis must look like as well as for evaluating the existing ones.

In chapter 4, I will present the experimental work that has been conducted so far, directly or indirectly aiming to investigate the availability and status of ignorance effects of superlative modifiers as compared to comparative modifiers (most of the time), and help us adjudicate between a semantic and a pragmatic

approach to those effects. More, new insights into this debate will be provided by the studies I carried out, which are reported in chapters 5 and 6. In correspondence with what I have been saying so far regarding the focus of the literature on modified numerals, these chapters zoom in on what is taken to be the main modal implication triggered by superlative modifiers (and class B modifiers in general), viz. a speaker-related epistemic effect. In chapter 7, however, I will question whether we should not also capture a related modal effect, that is, *speaker indifference*. To illustrate this effect, suppose once again that I utter (4).

(4) Magda ate at least four souvlakis.

If you do not know how much I enjoy keeping track of what my mom eats, you might also interpret (4) as implying that I don't care about the exact number n of souvlakis my mom had, or I do not think it is relevant to specify n, as long as $n \geq 4$. This interpretation is what I call speaker indifference and I will better define it in chapter 7. Speaker indifference has not been discussed in the existing literature on modified numerals, although it figures more prominently, for instance, in the literature on free choice items, see an example in (23) adapted from Kratzer & Shimoyama's (2002) work on the German indefinite *irgendein*.

(23) Hans: Irgendjemand hat angerufen.
Irgend-one has called

Maria: # Wer war es?
Who was it? (Kratzer & Shimoyama, 2002, p. 10)

As Kratzer & Shimoyama remark, in (23), *Hans conveys that he doesn't know or care about who called, or thinks the identity of the speaker is irrelevant* (p. 9), which makes Maria's question about the precise identity of the individual that called a rather pragmatically odd follow-up. Kratzer & Shimoyama characterize those modal speaker-oriented effects as ignorance and indifference effects, respectively. In chapter 7, I will elaborate on speaker indifference effects of modified numerals, where I am going to approach them both theoretically and experimentally.

In the remainder of the present chapter, I will be concerned with yet another type of effect that numeral modifiers can trigger, which is much less studied compared to ignorance inferences, but still more so compared to indifference effects, i.e., variation inferences.

2.3 Obviation of speaker ignorance: Variation effects

It has been observed that ignorance interpretations can disappear once numeral modifiers appear in certain embedding environments. For instance, if we embed the modified numeral phrase from (4) in the scope of the quantifier *everyone* in (24), we can get an interpretation that lacks ignorance, as (24) can be uttered by a speaker that is knowledgeable of the precise number of souvlakis everyone ate, given that no one ate fewer than four souvlakis.

(24) Everyone ate at least four souvlakis.

On such an interpretation, (24) usually gives rise to a so-called *variation* effect (Nouwen, 2015), where not everyone ate the same number of souvlakis. The resulting interpretation constitutes what Büring (2008) has called *authoritative reading* for examples where *at least* interacts with a universal modal, as in (25).

(25) (To win the eating contest,) Magda has to eat at least four souvlakis.
$\rightsquigarrow$ *the number of souvlakis Magda eats varies wrt the worlds in which Magda is eligible to win the contest* (variation implication)

In general, Nouwen (2015) observes that this reading becomes available through the interaction with a variety of operators, such as universal nominal and modal quantifiers as well as plurals and generics (see also Mayr, 2013), while according to Mayr's relevant work it is further concluded that the interaction with existential quantification does not trigger the effects in question. When numeral modifiers appear in the former embedding environments ignorance interpretations play second fiddle (cf. McNabb & Penka's (2014) experimental findings on interaction with universal modals), unless clearly supported or biased by the context. Most theories treat these two interpretations as a case of scope ambiguity (Büring, 2008; Cohen & Krifka, 2014; Coppock & Brochhagen, 2013b; Geurts & Nouwen, 2007; Kennedy, 2015; Nouwen, 2015): Ignorance arises when the superlative quantifier has wide scope with respect to the present operator, while when being in the scope of that operator the authoritative reading obtains. The aforementioned reading preference/saliency then might have to do with a well-known preference for surface scope readings as opposed to inverse scope readings (see, e.g., Anderson, 2004; Kurtzman & MacDonald, 1993; Tunstall, 1998 for experimental evidence).

That variation effects are available implications of numeral modifiers is suggested by the following contrast: Uttered by a knowledgeable speaker, (24) and (25), repeated below for convenience, are compatible with variation scenarios as in (26-a) and (27-a), respectively, while (24) is infelicitous in a situation in which every individual ate the same number of souvlakis, as in (26-b), and (25) is infelicitous in a scenario where the same number of souvlakis appears in each (goal-oriented) deontically accessible world, as in (27-b).

(24) Everyone ate at least four souvlakis.

(25) Magda has to eat at least four souvlakis.

(26) a. e1: 4 souvlakis
e2: 5 souvlakis
e3: 7 souvlakis
e4: 5 souvlakis

b. e1: 5 souvlakis
e2: 5 souvlakis
e3: 5 souvlakis
e4: 5 souvlakis
...

(27) a. w1: 9 souvlakis
w2: 4 souvlakis
w3: 6 souvlakis
w4: 6 souvlakis

b. w1: 5 souvlakis
w2: 5 souvlakis
w3: 5 souvlakis
w4: 5 souvlakis
...

e: individual entity w: deontically accessible world

In scenarios (26-a) and (27-a), which make (24) and (25) true and felicitous, respectively, we have a range of pairs of individuals and souvlakis, and of pairs of worlds and souvlakis, respectively. This is the result of the distribution of relevant numbers of souvlakis over individual people or deontically accessible worlds. The latter case is akin to similar effects observed with certain epistemic indefinites known as the *distribution requirement* (Kratzer & Shimoyama, 2002), illustrated by (28).

(28) Mary muss irgendeinen Arzt heiraten.
Mary has to irgend-a doctor marry
'Mary has to marry a doctor.'
$\rightsquigarrow$ *any doctor is a permitted option*
(adapted from Kratzer & Shimoyama, 2002, p. 12)

Because of the inference in question, (28) can be felicitously uttered in a situation like (29), but not like (30).

(29) w1: doc1
w2: doc2
w3: doc3
w4: doc4
...

(30) w1: doc2
w2: doc2
w3: doc2
w4: doc2
...

(31) w1: doc2
w2: doc3
w3: doc3
w4: doc2
...

Kratzer & Shimoyama (2002) further point out that the distribution requirement has it that there be an accessible world for *every* male doctor in the universe of discourse. Hence, assuming the same domain as in the case of (29), i.e., with the same number of male doctors, (28) does not felicitously describe (31) either. In other words, the infelicity is due to the fact that in (31) there are just two doctors, while the domain has more doctors. This effect is also known

as *total variation* and is actually the type of effect that embedded disjunction gives rise to resulting in a free choice interpretation (Kamp, 1973; Zimmermann, 2000, inter alia) as in (32) (see Nickel, 2011 and Crnič, Chemla & Fox, 2015 for corresponding inferences with disjunction embedded under generics and universal nominal quantifiers, respectively, also referred to as *distributive* inferences).

(32) Magda has to eat a mushroom souvlaki or a halloumi souvlaki.
$\rightsquigarrow$ *Magda is free to choose between a mushroom souvlaki and a halloumi souvlaki* (free choice inference)

As Nouwen (2015) remarks, the relevant effect that arises with superlative modified numerals in the scope of a universal modal is not a total variation effect but rather a weaker effect, known as *partial variation* or *modal variation*. Indeed, when not read with a speaker ignorance implication, (25) conveys that Magda does not eat the same number of souvlakis in every deontically accessible world and certainly lacks the total variation implication *for every number ≥4, there is a deontically accessible world where Magda eats that many souvlakis*. Nouwen (2015) illustrates this point by means of the following example:

Context: *Password policy:* For security reasons, the system will not accept passwords that are shorter than six characters. Moreover, it cannot handle passwords that are longer than ten characters.

(33) Passwords have to be at least six characters long.

(adapted from Nouwen, 2015, p. 249)

(33) is true and felicitous given the preceding **context**. If the variation effect of (33) were as strong as total variation (or free choice), (33) should not be felicitous, as it would imply that a password of *any* number of characters greater than 5 would be an acceptable one. Thus, the effect in question seems to merely say that there is no specific number n such that passwords need to be exactly n characters long, that is, there has to be at least two distinct numbers ≥ 6 (in the relevant domain) such that there is a password with as many characters.

In the following, I am turning to how the (partial) variation effects of modified numerals come about according to various theoretical accounts.

2.3.1 Accounts of variation effects of numeral modifiers

In contrast to the case of speaker ignorance effects, there are only few explicit accounts of variation effects of superlative modified numerals despite the fact that most of the existing accounts do derive the relevant asserted authoritative interpretation (Büring, 2008; Cohen & Krifka, 2014; Coppock & Brochhagen, 2013b; Geurts & Nouwen, 2007; Kennedy, 2015; Nouwen, 2015). Like speaker ignorance effects, variation effects have been tackled mostly from a pragmatic

perspective. Although the relevant accounts derive a partial variation effect such that at least two alternatives in the relevant domain are true for at least one individual entity e or world w, depending on the embedding environment (i.e., nominal or modal universal quantifier, respectively), they are divided as to the type of partial variation they derive: (i) There are those accounts where the output variation effect specifies which alternatives those may be and particularly includes the exhaustive interpretation of the minimum value compatible with the modified numeral phrase (Büring, 2008; Kennedy, 2015; Mayr, 2013; Schwarz, 2016a as well as Nouwen's (2015) overall proposal)[8], and (ii) those accounts where the partial variation implication they derive does not include any such specification (Coppock & Brochhagen, 2013b and Nouwen's (2015) anti-specificity-based (only) proposal). Notice the parallel to the two types of ignorance implications existing in the corresponding literature, i.e., *specific* vs. *underspecified ignorance*, identified earlier in section 2.2.3. Given this parallel, I will call the relevant inferences *specific variation* and *underspecified variation* inferences, respectively. As will become clear below, which type of inference we derive (either in the case of ignorance or of variation effects) has to do with the (type of) alternatives fed into the pragmatic mechanism we employ.

2.3.1.1 Accounts of *underspecified* partial variation effects

Coppock & Brochhagen's (2013b) account as well as the anti-specificity-based part of Nouwen's (2015) proposal belong to the class of accounts that derive *underspecified* partial variation effects. Although they differ in the derivation details, both accounts posit only exact values as alternatives to superlative numeral modifiers. In the following, I will exemplify this category with the latter proposal.

Remember that on Nouwen's (2015) proposal, *at least* is associated with an anti-specificity lexical requirement such that the domain of the cardinality of the set it expresses is not a singleton, i.e., there is no specific number or single option in that domain. Nouwen (2015) borrows (through Alonso-Ovalle & Menéndez-Benito's (2010) implementation) Kratzer & Shimoyama's (2002) reasoning in their analysis of free choice effects of the German modal indefinite *irgendein*, applied here to (25): Upon hearing (25), which asserts $\Box[4, ...)$ and has an anti-specificity presupposition, the listener wonders why the speaker did not use an (alternative) expression with a smaller, singleton domain resulting in a stronger statement, such as those in (34). She then concludes that the speaker implicates that for no singleton set of cardinalities $\{n : \ n \geq 4\}$ it is required that the cardinality (of the set) of souvlakis Magda eats is in that set (according to speaker's beliefs). Hence, the speaker implicates the negation of each of the alternatives in (34), that is, schematically, $\neg\Box[4]$, $\neg\Box[5]$, $\neg\Box[6]$, ...

[8] By *overall proposal*, I mean the combination of the anti-specificity-based reasoning and the Quantity-based reasoning Nouwen (2015) posits, recall from section 2.2.2.

(34) It is required that Magda eats exactly four souvlakis.
It is required that Magda eats exactly five souvlakis.
It is required that Magda eats exactly six souvlakis.
...

From assertion and implicatures, it follows that $\Diamond[4 \vee 5] \vee \Diamond[5 \vee 6] \vee \Diamond[4 \vee 6] \vee \ldots$. Consequently, there are two or more distinct numbers ≥ 4 such that Magda can eat that many souvlakis (and be eligible to win the eating contest). This is the underspecified partial variation reading we are after, which both Coppock & Brochhagen (2013b) and Nouwen (2015) derive drawing on Kratzer & Shimoyama's (2002) particular proposal involving *exact* alternatives, though they motivate those alternatives differently: The former through an alternative-introducing semantics and the latter via an anti-specificity presupposition.

2.3.1.2 Accounts of *specific* partial variation effects

In this category, we have the accounts that derive partial variation effects as scalar implicatures (secondary Quantity implicatures, Sauerland, 2004). Such are Büring's (2008), Kennedy's (2015), Mayr's (2013), Schwarz's (2016a) explicit proposals as well as the other half of Nouwen's (2015) proposal combined with the anti-specificity-based reasoning. Abstracting away from the details of these proposals, I will present what the main idea is that all of them share, as I did for the derivation of speaker ignorance effects as primary Quantity implicatures (see section 2.2.2). Although Mayr (2013), Nouwen (2015), and Schwarz (2016a) posit a more articulate set of alternatives, here too, I will only assume $[n]$ and $[n+1, \ldots)$ as the stronger alternatives to an *at least n* utterance, similarly to Büring (2008) and Kennedy (2015). Note that these two alternatives are anyway those that get to enter the scalar implicature reasoning in Mayr's (2013), Nouwen's (2015), and Schwarz's (2016a) accounts.

Let us consider again (25) and the context of the eating contest, repeated below as (35), and its stronger alternative propositions in (36).

(35) Magda has to eat at least four souvlakis.

(36) It is required that Magda eats exactly four souvlakis.
It is required that Magda eats more than four souvlakis.

If we execute the same Quantity-based recipe as in section 2.2.2, we get the primary implicatures in (37). Importantly, these implicatures are further strengthened to the secondary implicatures (according to Sauerland's (2004) distinction), in (38), such that the speaker believes that it is not the case that Magda has to eat exactly 4 souvlakis and she believes that it is not the case that Magda has to eat more than 4 souvlakis.

(37) $\neg\Box_{Bel}(\Box[4]), \neg\Box_{Bel}(\Box[5, \ldots)\ )$ *primary implications*

(38) $\Box_{Bel}(\neg\Box[4]), \Box_{Bel}(\neg\Box[5, \ldots)\ )$ *secondary implications*

Taking the implicatures in (38) together with the assertion $\Box[4,\ldots)$, actually together with the Quality implication $\Box_{Bel}(\Box[4,\ldots))$, it follows that $\Box_{Bel}\Diamond[4]\wedge \Box_{Bel}\Diamond[5,\ldots)$, i.e., according to the speaker, Magda can eat four souvlakis and she can eat more than 4 souvlakis and win the eating context. That is, there are at least two distinct numbers ≥ 4 such that Magda can eat that many souvlakis and win the contest, and one of these numbers is 4. This is how we arrive at the desired specific partial variation reading, which we can simplify to $\Diamond[4] \wedge \Diamond[5,\ldots)$, omitting the belief operator.

2.3.1.3 Variation effects of comparative modifiers

Before ending this section on variation effects, I would like to briefly turn our attention to comparative numeral modifiers. Of all the aforementioned accounts, only Mayr (2013) discusses variation effects with comparative modifiers, the rest implying that they are non-existing, like ignorance effects.

Mayr (2013) treats comparative modifiers in the same way as superlative modifiers with respect to variation inferences. He derives them as scalar implicatures via the pragmatic mechanism described in the previous subsection, that is, he derives the specific variation implication given below.

(39) Magda has to eat more than three souvlakis.
$\leadsto$ *Magda can eat exactly four and Magda can eat more than four souvlakis*

To conclude, Mayr's (2013) account predicts that variation implications arise with both types of numeral modifiers to the same extent.

On the other hand, Kennedy (2015) account does not derive variation implicatures (or scalar implicatures) with comparative modifiers because these modifiers do not have relevant (stronger) alternatives to be fed into the standard recipe. As a result, they cannot generate either primary or secondary Quantity implicatures. On Nouwen's (2015) or Coppock & Brochhagen's (2013b) account, variation effects do not arise with comparative modifiers because these modifiers are not associated with an anti-specificity presupposition or an alternative-introducing semantics, respectively, which are at the core of the pragmatic mechanism these accounts adopt in order to derive variation effects with superlative modifiers.[9]

In this section, we have been been dealing with the little investigated partial variation effects of modified numerals. The few existing theories diverge in (i)

[9]However, it is not clear how Coppock & Brochhagen (2013b) end up deriving variation effects *only* with superlative modifiers. They assume that Kratzer & Shimoyama's (2002) distribution requirement is calculated after existential closure has applied to the set of possibilities introduced by the superlative, outputting a singleton containing the union of those possibilities inside the scope of the present modal operator, i.e., $\{\Box\{P1 \cup P2 \cup P3 \cup \ldots\}\}$. This is exactly the denotation they assume for comparative modifiers, that is, of the respective comparative utterance. Hence, one would expect that, if the distribution requirement applies after existential closure, it should apply in the case of comparatives too, as they share the same (input-)denotation with superlatives at this stage of interpretation.

the type of variation effects they derive: underspecified vs. specific variation, and (ii) whether they derive such effects with both superlative and comparative numeral modifiers. The next chapter (chapter 3) aims to address and settle both issue (i) and (ii) providing novel data obtained by means of a series of experiments. Also, it will provide insight into the pragmatic mechanism associated with modified numerals. After this chapter, as already indicated, we will turn to the most well-investigated and known inferences of modified numerals, viz. speaker ignorance effects. In particular, we will deal with speaker ignorance effects in chapters 4, 5, and 6, where the theoretical background presented in the previous section (section 2.2) and the particular aspects we highlighted therein as well as the findings of the variation experiments in chapter 3 will prove very relevant and useful.

CHAPTER 3

Variation effects and their experimental investigation

3.1 Introduction

In the last part of chapter 2, we focused on the hitherto little-studied and empirically unexplored partial variation implications of modified numerals. We saw that the few explicit accounts of such effects differ in the way they capture those effects: specific variation vs. underspecified variation, which is a difference we also pointed out for ignorance inferences. We further observed that which type of inference we derive (either in the case of ignorance or of variation) has to do with the (type of) alternatives we feed into the relevant pragmatic mechanism. In this context, the present chapter sets out to investigate experimentally and provide novel data as to the availability, the strength, and the type of partial variation effects of modified numerals (see first set of questions in section 1.4). I do so by means of a series of experiments: experiments 1a, 1b and 1c test the likelihood and strength of variation effects with superlative modified numerals as well as with other types of modifiers, and experiments 2a and 2b examine which type of variation effects, underspecified or specific variation, is the right one and should be captured by a theoretical account.[1] My ultimate goal is to investigate the implicature mechanism and the alternatives involved in the generation of variation effects and, thus, to also adjudicate between the different existing theoretical proposals.

[1]Parts of experiments 1a and 2a have been reported in a paper that appeared in SALT 25 proceedings as Alexandropoulou et al. (2015) and part of experiment 1a has also been included in Alexandropoulou (2015) appearing in the proceedings of Sinn und Bedeutung 19.

3.2 Experimental method

All experiments presented in this chapter use the same method. Here, I will explicate this method, so that I do not repeat it for each individual experiment to be discussed. The method in question is a felicity judgement task where participants are asked to provide their judgement on a -3 to 3 Likert scale. Katsos & Bishop (2011) have shown that when asked to provide gradient, rather than binary, ratings for utterance-felicity, participants successfully distinguish among semantic entailments, logical contradictions, and pragmatically infelicitous statements. We chose to use a gradient modified Likert scale because we are interested in such an identification as regards variation effects of modified numerals. The particular type of scale we employed was specifically inspired by Cummins & Katsos (2010), who devised and used a -5 to 5 Likert scale in their experiments. In light of Katsos & Bishop's (2011) finding, Cummins & Katsos predicted that logical contradictions will be judged as completely infelicitous and score at the lower part of the scale, that items involving a logical entailment and being pragmatically felicitous would be judged as coherent, with scores at the upper part of the scale, while items that involve a pragmatic infelicity (e.g., implicature cancellation) would be judged as more coherent than the former, but yet less coherent than the latter, scoring at the middle of the scale. Indeed, Cummins & Katsos's (2010) particular response scale was utilized by participants as predicted.

Let me now illustrate the precise task we administered to our participants. We presented them with claims made by a researcher, followed by a question posed by someone interviewing the researcher. Participants were asked to judge to what extent the question made sense given the statement just made. Here are two illustrations of our task:

(1) **Researcher:** Some of my participants were left-handed.
Interviewer: Were all of your participants left-handed?
Does the interviewer's question make sense?

(2) **Researcher:** Some of my participants were left-handed.
Interviewer: How did you find out that not all of your participants were left-handed?
Does the interviewer's question make sense?

Participants had to indicate to which extent they thought the interviewer's question made sense given the claim just made by the researcher by picking a score on the -3 to 3 Likert scale. If participants calculate the (scalar) implicature in the researcher's claim (i.e., that not all of the researcher's participants were left-handed), then they will judge the interviewer's question in (1) as relatively infelicitous, as they will assume that the interviewer should already know the answer to the question she is asking. On the other hand, in (2), calculating the implicature will lead to a higher felicity score, since the question takes for granted that the implicature has been calculated by the interviewer. In both

variants of the task, the interviewer's question targeting the scalar implicature (in one way or another), an optionally drawn inference, is expected to receive felicity ratings higher than the corresponding bad controls (where the interviewer's question makes no sense whatsoever given the researcher's claim) and lower than the corresponding good controls (where the interviewer's question makes complete sense given the researcher's claim). By using (either variant of) this method, we can measure the likelihood that a certain inference is drawn. Experiments 1a, 1b, and 1c use the paradigm in (1), while experiments 2a and 2b use that illustrated by (2), without there being a deeper reason why we used two variants of the task.

3.3 Variation experiments 1a, 1b and 1c

Experiments 1a, 1b, and 1c were conducted in Dutch and aim to examine the likelihood and strength of partial variation effects with superlative numeral modifiers as well as with other types of modifiers. All three experiments target variation effects in embedding environments created by a universal nominal quantifier. Experiment 1a tests lower-bound numeral modifiers, experiment 1b tests upper-bound numeral modifiers, and experiment 1c constitutes a follow-up to experiment 1a.

3.3.1 Variation experiment 1a

3.3.1.1 Design & Material

The experimental conditions involved statements uttered by the researcher with a lower-bound modified numeral embedded in the scope of the universal nominal quantifier *elk* 'each'. As specified in the instructions, the researcher's statement reported on findings of a recent successful research s/he was involved in. The interviewer's question asked about more details regarding the researcher's statement in a way that went against the variation effect of the researcher's statement. We manipulated the form of the numeral modifier (NM) in the researcher's statement: *superlative* (*minstens* 'at least'), *comparative* (*meer dan* 'more than'), *disjunctive* (*n of meer* '*n* or more'). The experiment also included another manipulation regarding the type of question the interviewer asks, which was part of a separate experiment. The relevant factors were manipulated in a 3×2 design. The latter manipulation will not be discussed in the present thesis, as it has nothing to do with the questions tackled here. In (3) you are given an example of an experimental item and its English translation in (4) in all three NM conditions.

(3) ***Onderzoeker:*** *Tijdens het evenement werd elke straat door*
researcher during the event was each street by
{minstens zes / meer dan zes / zes of meer} *agenten beveiligd.*
{at least six / more than six / six or more} policemen guarded

Interviewer: *Werden ze allemaal door evenveel agenten beveiligd?*
interviewer were they all by as many policemen guarded

(4) **Researcher:** During the event every street was guarded by
{at least six / more than six / six or more} policemen.
Interviewer: Were they all guarded by the same number of policemen?

As you notice, the interviewer's question targets the (partial) variation inference of the researcher's statement such that *it is not the case that every street was guarded by the same number of policemen.* Taking into account this inference, the interviewer's question feels as an unreasonable and infelicitous follow-up to the researcher's claim, as it asks whether the exact opposite is the case. The use of the researcher-interviewer setup serves to weaken a potential ignorance interpretation of the researcher's statement (i.e., *every street was guarded by n policemen, and as far as the speaker knows n could be 6 or more*). The assumption is that the researcher (who is reporting on findings of a recent successful research s/he was involved in, as stated in the instructions) is most likely to be an authority on her/his own findings.

The experimental task included six test items and six filler items. It further included thirteen good control items and four bad control items (number of stimuli = 29).[2] In the good control items, the interviewer asks a relatively neutral follow-up question, which does not contradict the researcher's statement. In the bad controls, the interviewer's question prompts a proposition that contradicts the researcher's statement. Examples of good and bad controls are given in (5) and (6), respectively (without glosses).

(5) ***Onderzoeker:*** *Elke boer in Drenthe heeft zes of meer koeien.*
'**Researcher:** Each farmer in Drenthe has six or more cows.'

Interviewer: *Hoe ben je daar achtergekomen?*
'**Interviewer:** How did you find that out?'

[2]It was expected that the test items would be judged as bad/infelicitous due to the availability of variation effects, so more good controls were added to offset the asymmetry.

(6) ***Onderzoeker:*** *Het CBS voorspelde dat de economie in 2013 zou krimpen.*
'**Researcher:** CBS predicted that the economy would shrink in 2013.'

Interviewer: *Waaraan dacht het CBS dat die groei te danken viel?*
'**Interviewer:** What did CBS think was the cause of this growth?'

As you can tell, the interviewer's question in (5) is a reasonable and natural follow-up to the researcher's claim, while the interviewer's question in (6) has the presupposition that there is growth, which is false given the researcher's statement, hence, it is severely infelicitous. Five of the good control items involved the universal quantifier *elk* ('each') and a modified numeral in the researcher's statement, as in (5), while the other eight contained completely different configurations, like the bad controls (see (6)) and the filler items. Lastly, the experiment also included four practice items, which had the form of bad and good controls, presented above; that is, two of them were to be rated high on the felicity scale and the other two were to receive low felicity scores.

The test items appeared in six conditions (3 NMs × 2 types of question). They were rotated through six lists, so that each participant saw each item only in one experimental condition. Control items were the same across lists.

3.3.1.2 Predictions

First, it is expected that our bad controls will be judged as highly infelicitous receiving only scores from the bottom of the scale, while the good controls consisting of neutral and felicitous researcher-interviewer interactions should obtain scores from the upper part of the scale.

More importantly, if partial variation is an implicature that superlative modified numerals give rise to in embedding contexts, we expect that participants that calculate this implicature in the researcher's statement will judge the interviewer's question as infelicitous. If this is a strong/robust implication, we expect the infelicity of the respective test items to resemble that of bad controls, where the interviewer's question prompts a proposition that contradicts the researcher's assertion, and to be clearly different from the condition of good controls. If it is a non-obligatory implication, which might or might not be calculated by all participants, we expect to find the superlative condition scoring between bad controls and good controls.

Furthermore, as to the comparative NM condition, if comparative modified numerals trigger partial variation effects to the same extent as superlative modified numerals when in the nuclear scope of a universal quantifier, as predicted by Mayr's (2013) account, comparative items should be judged in a similar way to superlative items. On the other hand, if partial variation is not an available implication of comparative modified numerals, as implicitly predicted by a number of accounts, such as Coppock & Brochhagen (2013b); Kennedy (2015), and Nouwen (2015), we should find a contrast between comparative and superlative items. Comparative items should be judged as more felicitous, because the interviewer in these items does not calculate a variation implicature in the

researcher's statement and, thus, does not (already) know the answer to the question s/he is posing, which is the case in the superlative condition. Hence, the question s/he asks is expected to be perceived as neutral and plausible, as in the case of good control items.

Concerning the disjunctive NM condition, we expect a similar behavior to that of superlative items. *N of meer* ('*n* or more') has been characterized as a class B modifier by Nouwen (2010), like superlative modifiers, and is also associated with an anti-specificity presupposition (Nouwen, 2015). At the same time, *n of meer*, coming in a disjunctive form and being the disjunctive equivalent of the lower-bound superlative modifier, perfectly spells out Büring's (2008) proposed disjunctive representation of the lower-bound superlative modifier or the Horn-scale-based pragmatic alternatives Kennedy (2015) posits for the lower-bound superlative modifier. It could also be seen as summarizing the multiple-alternative-based denotation of the lower-bound superlative modifier on Coppock & Brochhagen's (2013b) account. Thus, Nouwen (2015), Coppock & Brochhagen (2013b), as well as Büring (2008) and Kennedy (2015) predict that variation effects will be triggered in the researcher's statements in the disjunctive condition to the same extent as in the superlative condition, with the respective dialogues resulting in the same degree of infelicity.

Table 3.1 lists the aforementioned predictions according to each theoretical account.

Theoretical accounts	**Predictions re NM conditions**
Büring (2008), Kennedy (2015)	comparative > superlative disjunctive = superlative
Coppock & Brochhagen (2013b)	comparative > superlative disjunctive = superlative
Mayr (2013)	comparative = superlative
Nouwen (2015)	comparative > superlative disjunctive = superlative

Table 3.1: Predictions of theoretical accounts as to experiment 1a's conditions involving different types of numeral modifiers. Symbols > and = stand for higher and comparable felicity ratings, respectively.

3.3.1.3 Participants

97 people participated in experiment 1a out of whom only 68 were native speakers of Dutch who filled in the entire questionnaire without making any mistakes on the practice items. We used only the data from these 68 participants (39 female, mean age: 38, age range: 20–68) in subsequent analyses. All of the participants volunteered in filling in the questionnaire and were naive as to the purpose of the study.

3.3.1.4 Procedure

Participants were presented with items like (3)-(6) and were asked to judge how well the interviewer understood the researcher's claim. They did so on a -3 to 3 scale, where -3 is "the claim is not understood" and 3 is "the claim is understood".[3] Participants first read the instructions and then they would move on to the practice block to familiarize themselves with the experimental task. The actual experiment started right after.

The lists of the experiment were created using the free on-line survey tool SurveyMonkey (`www.surveymonkey.com`), where links to each list were generated. The lists were randomly distributed and filled in mostly on-line. Each and every trial of the experiment block was presented on a single screen. After the on-line collection of data, a few more data were obtained in a pen-and-paper fashion in order to balance the number of observations across lists.

3.3.1.5 Results

As said, the data of sixty-eight participants were kept for statistical analysis, the final number of observations being $N = 1564$. Part of the data of a bad control item and of a good control item was not included in the analysis because of a typo. The data of these items were not discarded altogether, because the typo was noticed and corrected in time, and the relevant link to the respective list of the experiment was made available again.

Participants' responses were ordered categorical, thus the collected data were analyzed with ordered probit models using the `ordinal` package (Christensen, 2013) in `R`. We conducted two analyses: (i) an analysis including the type of NM as predictor, with the reference level being the superlative NM condition; (ii) an analysis where the reference level of the previous analysis is contrasted with the conditions of good and bad controls, thus, by including one fixed effect again. The analyses also included intercept and slope random effects for participants and items. We applied backward model selection for random effects (Barr, Levy, Scheepers & Tily, 2013) and in the following I will be presenting the output of the model with the maximal random-effect structure that converged and had the best fit.

[3]The respective Dutch sentences were *de bewering is niet begrepen* and *de bewering is begrepen*.

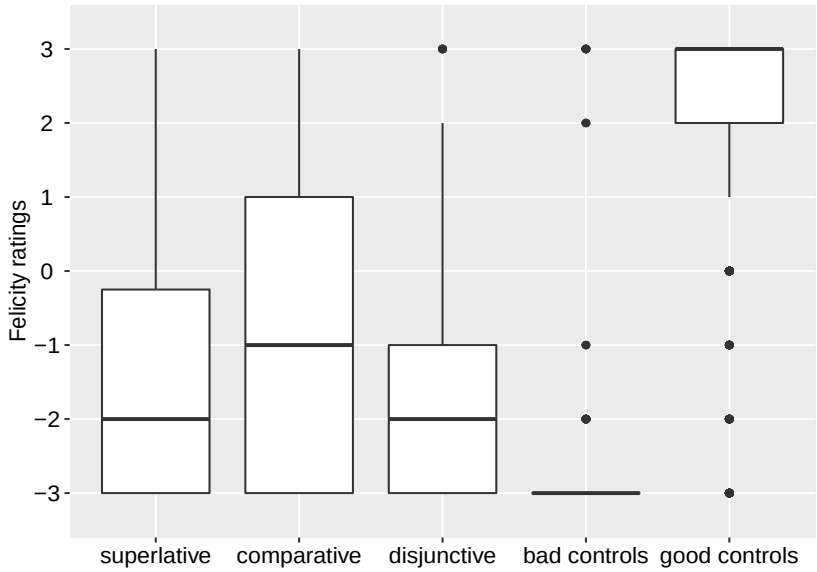

Figure 3.1: Box plots of felicity ratings per NM and control condition.

The box plots in figure 3.1 present the felicity ratings per NM and control condition, see respective item examples in (3)-(4) (NM conditions), in (5) (good controls) and (6) (bad controls). As the box plots reveal, the bad and good control items were judged as expected, with the former scoring very low on the felicity scale ($mean = -2.792$, $median = -3$), and the latter receiving high felicity ratings ($mean = 2.128$, $median = 3$). As is also evident, the superlative NM items were generally judged as being infelicitous ($mean = -1.485$, $median = -2$), but were found to be (highly) significantly more felicitous than bad controls ($\beta = 1.895$, $SE = .416$, $z = 4.557$, $p < .0001$), and reliably less felicitous than the good controls ($\beta = -2.409$, $SE = .276$, $z = -8.726$, $p < .0001$), according to the results of the second analysis (see (ii) on previous page).

Moreover, we observe that comparative items were judged to be relatively infelicitous too, though receiving higher scores than the superlative items ($mean = -.632$, $median = -1$). The disjunctive NM condition exhibits a similar behavior to the superlative condition, receiving low(er) scores overall ($mean = -1.735$, $mean = -2$). The analysis we conducted on the different types of NMs (analysis (i)) showed that the appearing difference in felicity ratings between the superlative and the comparative condition is significant ($\beta = .583$, $SE = .203$, $z = 2.892$, $p < .01$), while there is no significant difference between the superlative and the disjunctive condition ($\beta = -.241$, $SE = .206$, $z = -1.168$, $p = .243$).

3.3.1.6 Discussion

As has already been specified, the interviewer's question in the superlative condition would be a felicitous and plausible follow-up to the corresponding statement by the researcher, if it was not perceived as going against the relevant variation inference, that is, if such an inference was not triggered by the researcher's statement. The fact that superlative items were judged as significantly less felicitous than the good control items indicates that the researcher's statements in the superlative items gave rise to variation effects. Moreover, superlative items were found to score significantly higher than the bad control items, which shows that the variation effects triggered in the researcher's statements are non-obligatory inferences, thereby confirming their pragmatic status.

Comparative items were also judged as infelicitous overall (cf. the difference from the good control items) but importantly they received higher felicity ratings than superlative items. Based on that, we conclude that variation effects become available with comparative utterances too, but they are less robust than with superlative utterances. This finding goes against the predictions of all accounts in table 3.1. Although Coppock & Brochhagen (2013b); Kennedy (2015), and Nouwen (2015) predict a difference between the comparative and the superlative NM condition, they further predict that the interviewer's questions in comparative items will be judged as felicitous and plausible followups, similarly to the good control items, which is clearly contrary to fact. In contrast, Mayr (2013) predicts that variation implications arise to the same extent with comparative and superlative utterances. This, too, is falsified by our findings.

Lastly, disjunctive items generally received low felicity scores, and somewhat lower than superlative items. This difference was not found to be significant, which is in line with the position that disjunctive and superlative modified numerals (which are both class B modified numerals) trigger variation implications to the same degree, as predicted under Büring's (2008), Coppock & Brochhagen's (2013b), Kennedy's (2015), and Nouwen's (2015) accounts.

This strength difference in variation effects exhibited between comparative modified numerals, on the one hand, and superlative and disjunctive modified numerals, on the other, is also attested among their upper-bound counterparts. The next section presents the relevant results.

3.3.2 Variation experiment 1b

3.3.2.1 Design & Material

Experiment 1b is very similar to Experiment 1a. First, we tried to test the same experimental items and modify them if needed. As a result, three of the items were kept intact, one was slightly modified, one was modified more, and another one was changed completely. A further difference is that Experiment 1b did not

include two manipulations but only that of the NM type. The superlative NM we tested was *maximaal* ('maximally'), from the Latin *maximum*, which is the neuter of the adjective *maximus*, the superlative form of *magnus* 'great'. The comparative and disjunctive modifiers we tested were *minder dan* ('less/fewer than') and *n of minder* ('*n* or less/fewer'), respectively (see part A.2 in Appendix A for the list of all test items).

The task included four practice items, six experimental items and four filler items, as well as the thirteen good control items from experiment 1a and three of its bad control items (number of stimuli = 30). Experimental items appeared in all three NM conditions. They were rotated through six lists, so that each participant saw each item only in one experimental condition. Control items were the same across lists.

3.3.2.2 Participants

Seventy people participated in experiment 1b, most of whom were students at the BA introductory course *Psychology of Language* at Utrecht University. Only 58 of all participants were native speakers of Dutch (only) who understood the task completely and filled in the entire questionnaire without making any mistakes on the practice items. We used only the data from these 58 participants (45 female, 1 without gender info, age mean: 19.1, age range: 18–28) for statistical analysis. All participants volunteered in filling in the questionnaire and were naive as to the purpose of the study.

3.3.2.3 Procedure

Contrary to experiment 1a, this experiment was in its entirety a pen-and-paper experiment. The procedure was as in experiment 1a: Participants first filled in a consent form, then they read the instructions and completed the practice block. The experimental block started right after.

3.3.2.4 Results & Discussion

Part of the data from the fifty-eight participants was excluded from the statistical analysis. These were observations of a bad control item whose mean was higher than expected, i.e., bigger than −1, and of two good control items whose mean was lower than expected, i.e., smaller than 1. The final number of observations was $N = 1098$. One statistical analysis was conducted, where the superlative condition was compared to the other two NM conditions as well as to the two control conditions.

The picture that arose is very similar to that in experiment 1a. The bad and good control conditions scored as expected (see figure 3.2), and the superlative condition was generally judged as infelicitous (*mean* = −.767, *median* = −1), scoring (highly) significantly higher than the bad control condition ($\beta = 2.295$, $SE = .435$, $z = 5.282$, $p < .0001$) and (highly) significantly lower than

the good control condition (β = −2.233, SE = .232, z = −9.611, p < .0001). Moreover, *maximaal* items were found to be significantly less felicitous than *minder dan* items (β = −.326, SE = .144, z = −2.267, p < .05), while their difference from the *n of minder* items was just significant (β = .298, SE = .147, z = 2.034, p = .042). Consequently, the relevant findings of both experiment 1a and 1b demonstrate that it is generalizable that comparative modified numerals embedded in the scope of a universal nominal quantifier trigger (partial) variation implicatures of a less robust nature compared to class B modified numerals.

In this experiment, we also found a difference between the superlative and the disjunctive NM conditions, suggesting that variation implicatures are even more robust with the disjunctive NM *n of minder*. This would go against the prediction arising from Büring's (2008), Coppock & Brochhagen's (2013b), Kennedy's (2015), and Nouwen's (2015) accounts that the two class B NMs give rise to variation effects of the same strength.

Finally, the felicity ratings of the NM conditions of experiment 1b are overall higher than the corresponding ratings in experiment 1a, which could be taken to imply that variation inferences are more robustly available with lower-bound modified numerals. Here, we are mainly interested in the relative differences and patterns within experiments and I leave the comparison between lower- and upper-bound numeral modifiers to further research. However, I should perhaps already note that for now we should avoid a direct comparison of experiments 1a and 1b, as only half of the experimental items were identical.

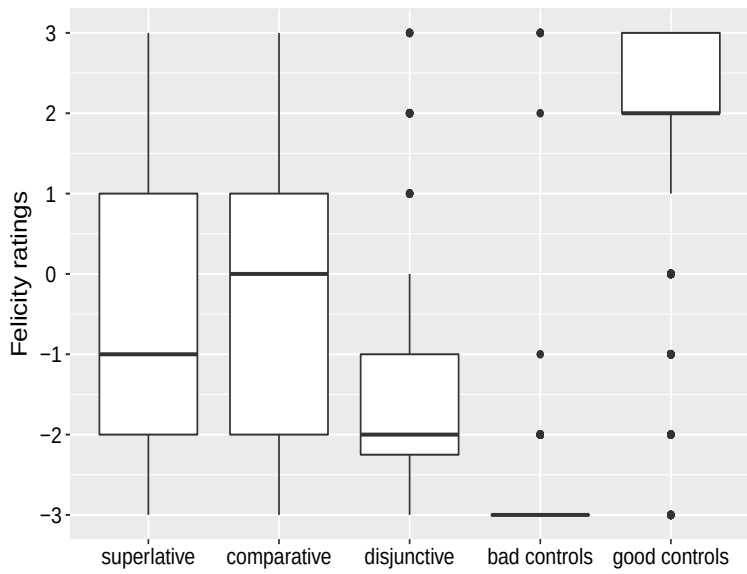

Figure 3.2: Box plots of felicity ratings per NM and control condition in experiment 1b.

There is a possible concern about experiments 1a and 1b and the findings thereof. As the experimental items in these experiments did not form minimal pairs with the control items, one could argue that there may be other causes for the relevant differences attested. Thus, we ran another version of experiment 1a in order to take care of this issue. In this follow-up experiment 1c, the bad and good control items, too, included modified numerals in the researcher's statements, and only differed from the experimental items in the interviewer's question. I am briefly reporting on this follow-up experiment in the next section.

3.3.3 Variation experiment 1c

3.3.3.1 Design & Material

In experiment 1c, we manipulated (i) the form of the NM: *superlative* (*minstens* 'at least') vs. *comparative* (*meer dan* 'more than'), and (ii) the question of the interviewer: *target/test*, *bad control*, *good control*. That is, the two factors were manipulated in a 2×3 design. Below you see an example of an item in all six conditions:

(7) **Onderzoeker:** *Tijdens het evenement werd elke straat door* $\left\{\begin{matrix}\text{minstens}\\ \text{meer dan}\end{matrix}\right\}$ *zes agenten beveiligd.*

Researcher: 'During the event each street was guarded by $\left\{\begin{matrix}\text{at least}\\ \text{more than}\end{matrix}\right\}$ six policemen.'

Test item **Interviewer:** *Werden ze allemaal door evenveel agenten beveiligd?*
Interviewer: 'Were they all guarded by the same number of policemen?'

Bad control **Interviewer:** *Was er geen straat met precies vier agenten?*
Interviewer: 'Was there no street with exactly four policemen?'

Good control **Interviewer:** *Hoe weet je dat?*
Interviewer: 'How do you know that?'

The test item is identical to that in experiment 1a. In bad controls, the interviewer's question prompted a proposition that contradicted the asserted meaning of the researcher's utterance, which for (7) is *for every street the number n of guarding policemen during the event is* ≥ 6. Thus, by asking such a question the interviewer shows that s/he has not understood the researcher's claim. In good controls, the interviewer's question was a neutral and plausible follow-up to the researcher's statement.

We added six more items (12 in total, see the whole list of experimental items in part A.3 of Appendix A) and included twenty-six fillers (total number of stimuli: 38). All items were rotated through lists so that each participant saw each item only in one condition.

3.3.3.2 Participants & Procedure

The experiment was created in Ibex and was hosted on Ibex farm (Drummond, 2007). We collected on-line judgements from thirty native speakers of Dutch (25 female, age range: 18–54, mean age: 23.033), recruited from the participant database of the UiL OTS lab. The task was the exact same as in the previous experiments.

3.3.3.3 Results & Discussion

As in experiment 1a, the superlative test items were judged as infelicitous in general ($mean = -.733$, $median = -1$), and significantly less felicitous than the good controls ($mean = 1.383$, $median = 1$, $\beta = -2.042$, $SE = .384$, $z = -5.319$, $p < .001$), but significantly more felicitous than the bad controls ($mean = -2.017$, $median = -3$, $\beta = 1.870$, $SE = .676$, $z = 2.769$, $p < .001$). Without doubt this time, the differences in felicity judgements indicate that the researcher's statements with superlative modified numerals give rise to non-obligatory variation implications of pragmatic nature.

Interestingly, we did not replicate the difference between the superlative and the comparative test items found in experiment 1a (or experiment 1b) nor did we find any significant interaction with either control condition. Comparative test items scored roughly as high as the superlative ones ($mean = -.817$, $median = -1$). This result is compatible with Mayr's (2013) prediction that comparative and superlative modified numerals give rise to variation effects to the same extent, and against the rest of the accounts that predict no variation implication for comparatives.

While failing to detect a difference in felicity ratings between the two types of NMs might be due to statistical power, we did observe a difference in the response times in the test condition. Running an Ibex experiment gave us the opportunity to also collect response times per condition, that is, for each trial, we measured the time from the appearance of the relevant item on a single screen until the participant made their judgment by clicking on a number from the response scale. Although this is an internet-based experiment, it was conducted in a controlled environment within a booth, so the obtained response times are highly reliable. Figure 3.3 presents the mean raw response times per NM per test/control condition, see respective item examples in (7). The linear mixed-effects regression analysis we conducted on the (log-transformed) response times revealed that (only) the difference we see between superlative and comparative test items is significant (main effect of NM in test condition: $\beta = .147$, $SE = .070$, $t = 2.099$, $p = .049$). This means that, although these two conditions ended up receiving roughly the same felicity scores, participants spend more time considering the comparative test items. Given that the output of the judgement process is similar, this result suggests that the difference has to do with the very process of judgement, revealing that this process is more effortful in the comparative test items. But what could be effortful about it?

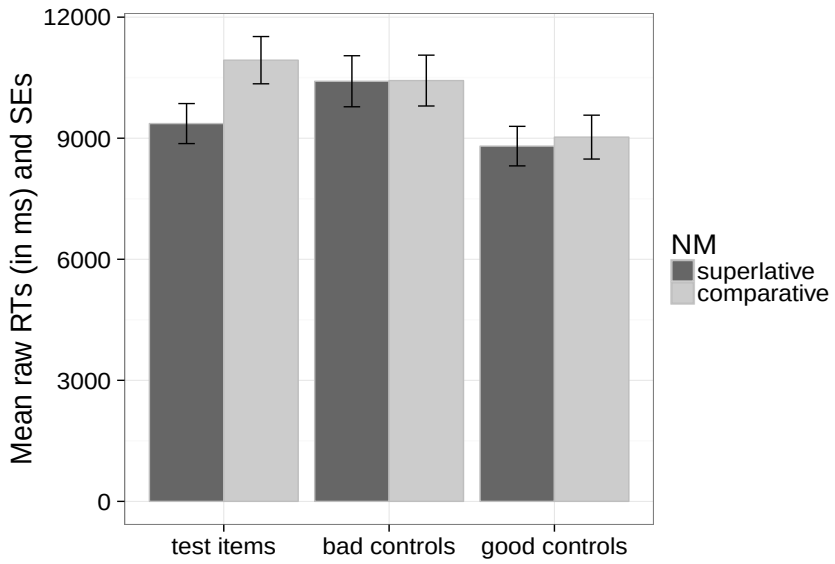

Figure 3.3: Response times per NM per test/control condition.

It could be that although participants derive the variation interpretation in the researcher's statement to the same extent in the superlative and the comparative condition, they are possibly less willing to do so in the comparative NM condition. Perhaps they consider that using the comparative modifier is pragmatically less appropriate in the particular environment compared to the superlative modifier or, in other words, that *minstens* ('at least') is a better cue to a variation implicature than *meer dan* ('more than'), resulting in a Manner-based reasoning.

Another possibility is that the extra effort attested in the comparative test items has to do with the process of accessing and calculating the particular variation inference in the researcher's statements. That is, it might suggest that the variation implications of comparative modified numerals are of a weaker, or just different, nature compared to superlative modified numerals, and more work and effort is needed to access such implicatures. Or relatedly, being less robustly associated with the comparative modifier, the variation implicature might be in general derived in the comparative items after participants are reminded to do so by reading the interviewer's question that specifically targets the variation interpretation of the researcher's statement.

Whichever scenario we take, we can conclude that the pragmatic reasoning that takes place in comparative test items is different from that in superlative test items.

3.3.4 Conclusion

To conclude the part of the current study concerning the likelihood and strength of variation effects, our overall results reveal that those effects are available with superlative modified numerals and they are non-obligatory implications, suggesting that deriving them via a pragmatic mechanism is the right way to go, in line with the relevant existing literature. The same seems to hold for variation effects with disjunctive modified numerals, another case of class B modified numerals.

On the other hand, although we find that variation inferences are available with comparative modified numerals too, in support of Mayr (2013), we also have an indication that these effects become less robustly available with comparatives evincing a weaker pragmatic nature as compared to superlative modifiers (or class B modifiers). This follows from the findings of both experiments 1a and 1b, but not from experiment 1c. Although we have not been able to replicate the relevant difference in felicity ratings in experiment 1c, this experiment did reveal a reliable difference in response times, which is in line with the scenario that variation implicatures are less robust with comparative modifiers. This is something that will further be addressed in the experiments in the remainder of the chapter, to which I turn next. The following experiments examine the partial variation effects in further depth. In particular, they seek to shed light on the type of partial variation modified numerals give rise: underspecified vs. specific partial variation.

3.4 Variation experiments 2a and 2b

Experiments 2a and 2b belong to the same study, also conducted in Dutch, and aim to find out which type of partial variation effects superlative modified numerals give rise to: *underspecified* vs. *specific* variation.[4] Experiment 2a tests the lower-bound superlative modifier in the scope of a universal nominal quantifier (as in experiments 1a and 1c), while experiment 2b tests the lower-bound superlative modifier in the scope of a universal modal quantifier. Below you see an example of each configuration and the relevant types of variation inferences.

(8) Everyone's mom ate at least four souvlakis.

$\rightsquigarrow$ *It is not the case that everyone's mom ate the same number of souvlakis* (underspecified variation)

$\rightsquigarrow$ *Some(one's) mom ate exactly four souvlakis and some(one's) mom ate more than four souvlakis* (specific variation)

(9) Magda has to eat at least four souvlakis.

$\rightsquigarrow$ *It is not the case that Magda eats the same number of souvlakis in*

[4]The design of the experiments I report in this section was joint work with Jakub Dotlačil, Yaron McNabb, and Rick Nouwen. I thank Jakub Dotlačil for obtaining the data and all of them for letting me report those data here.

all deontically accessible worlds (underspecified variation)
$\rightsquigarrow$ *Magda may eat exactly four souvlakis and Magda may eat more than four souvlakis* (specific variation)

As you can observe, the two types of variation differ in that specific variation is stronger in the sense that it additionally requires that the value 4 is true of (at least) one's mom in the nominal quantifier case and of (at least) one deontically accessible world in the modal quantifier case. This aspect will be used and targeted by experiments 2a and 2b.

Moreover, as explained in section 2.3.1, each type of partial variation stems from a pragmatic mechanism that employs a certain type of set of alternatives: underspecified variation comes about via reasoning about exact, maximally informative, alternatives (see (10-a) and (10-b), schematically: $\forall[4]$, $\forall[5]$, $\forall[6]$, ..., and $\Box[4]$, $\Box[5]$, $\Box[6]$, ..., respectively), while specific variation is derived via scalar implicature reasoning about binary alternatives (see (11-a) and (11-b), schematically: $\forall[4]$, $\forall[5,\ldots)$ and $\Box[4]$, $\Box[5,\ldots)$, respectively). Hence, by looking into the types of variation inferences, we moreover aim to draw conclusions as to the type of alternatives to be associated with superlative numeral modifiers.

(10) a. Everyone's mom ate exactly four souvlakis.
Everyone's mom ate exactly five souvlakis.
Everyone's mom ate exactly six souvlakis.
...
b. Magda has to eat exactly four souvlakis.
Magda has to eat exactly five souvlakis.
Magda has to eat exactly six souvlakis.
...

(11) a. Everyone's mom ate exactly four souvlakis.
Everyone's mom ate more than four souvlakis.
b. Magda has to eat exactly four souvlakis.
Magda has to eat more than four souvlakis.

Finally, the present experiment also tests the comparative lower-bound counterpart seeking to put into further investigation the contrast attested between the two types of NMs in experiments 1a and 1b.

3.4.1 Variation experiment 2a

3.4.1.1 Design & Material

The present experiment uses the researcher-interviewer setup, too, but employs the second paradigm described in section 3.2. We manipulated (i) the form of the numeral modifier (NM) in the researcher's statement, as in experiments 1a, 1b, and 1c: *superlative* (*minstens* 'at least') vs. *comparative* (*meer dan* 'more than'), and (ii) the type of embedding nominal operator, i.e., *alle* 'all' vs. *enkel*

'some', in combination with the interviewer's question. In a previous version of this experiment we ran, we included both *alle* ('all') and *elk* ('each') universal quantifiers as test conditions. As predicted by Mayr (2013) and Nouwen (2015) (i.e., that interaction with universal—vs. existential—quantification in general gives rise to variation), no differences were attested between the two test conditions in variation implicature rates (that is, in how the test conditions differed from control items). Hence, we randomly decided to continue testing with one of the two universal nominal quantifiers.

The above factors were manipulated in a 2×3 design.[5] In the following, I am presenting an experimental item in all six resulting conditions. First, by (12) I present four of the six conditions of the experiment, where the embedding operator is *alle* ('all').

(12) **Original item in Dutch**

Onderzoeker: *Tijdens het evenement werden alle straten door*
researcher during the event was all streets by
{minstens vier / meer dan drie} *agenten beveiligd.*
{at least four / more than three} policemen guarded

Interviewer: *Hoe ben je er achtergekomen dat er een*
interviewer how are you there found out.PTCP that there a
straat met {vier / vijf} *agenten is?*
street with {four / five} policemen is

Translated item in English

Researcher: During the event all streets were guarded by {at least four / more than three} policemen.

Interviewer: How did you find out that there was a street with {four / five} policemen?

In the above conditions, the researcher makes a statement as in experiments 1a-c, found to trigger variation effects. The underspecified variation interpretation has it that there is no specific number m such that all streets have m-many guarding policemen during the event. The specific variation interpretation additionally requires that there was (at least) one street that had (exactly)

[5]Experiment 2a also included a condition with disjunction testing variation effects, which was part of a separate study, not relevant to the purposes of the current study.

four guarding policemen. In this particular paradigm, the interviewer's question takes for granted—apparently given the information just conveyed by the researcher—that there was a street with n-many guarding policemen. As is obvious, when in the interviewer's question n is *four*, if the researcher's statement conveys the specific variation implication, the presupposed information in the interviewer's question is justified. If the researcher's statement conveys the underspecified variation, then the interviewer's presupposition is not satisfied and participants have to accommodate the presupposed information, which is compatible with the asserted meaning conveyed by the researcher (the number of guarding policemen all streets have is ≥4). From now on, I will call this condition, where the minimum value n is targeted by the interviewer's question, the *minimum value* condition.

Let us see what is supposed to happen regarding the condition where the number n in the interviewer's question is *five*, henceforth the *higher value* condition (5 is *higher* than the minimum value compatible with the modified numeral in the researcher's statement). The interviewer's question presupposes that there is some street with five guarding policemen. This information is not given by an inference associated with the researcher's statement, either on an underspecified or on a specific variation reading. Hence, being compatible with the asserted information in the researcher's statement, this information has to be accommodated in either case. For instance, participants can accommodate this information making the following reasoning: some streets are guarded by more than four policemen is compatible with some street having five policemen. The general idea behind this paradigm is that presenting a presupposition trigger without support for the presupposition will be less felicitous than when providing support for the presupposition. Although I will discuss the predictions in depth in the next section, I should already say that the predictions with respect to the *minimum value* and *higher value* conditions are different depending on the type of (partial) variation we assume.

Next, (13) presents the remaining two conditions with *enkele* ('some') as the embedding operator. Recall from chapter 2 (section 2.3), interaction with existential quantification does not trigger variation effects.

(13) **Original item in Dutch**

Onderzoeker: *Tijdens het evenement werden enkele straten door*
researcher during the event were some streets by
{minstens vier / meer dan drie} *agenten beveiligd.*
{at least four / more than three} policemen guarded

Interviewer: *Hoe ben je achtergekomen dat er geen straat*
interviewer how are you found out.PTCP that there no street
met meer dan vier agenten was?
with more than four policemen was

Translated item in English

Researcher: During the event some streets were guarded by $\left\{\begin{array}{c}\text{at least four}\\ \text{more than three}\end{array}\right\}$ policemen.
Interviewer: How did you find out that there was no street with more than four policemen?

Items with *enkel(e)* such as (13) served as bad control items and will be referred to as *some* bad controls. In these items, the interviewer's question presupposes that no street was guarded by more than four policemen. This information would be made available by the researcher's statement if we computed a scalar implicature on the basis of the scale $<$ *more than* n, *more than* $n+1$, $\ldots >$ for the comparative NM and $<$ *at least* n, *at least* $n+1$, $\ldots >$ for the superlative NM. However, as already said, such an implicature is not available when either comparative or superlative modified numerals appear in the scope of an existential nominal or modal quantifier (Fox & Hackl, 2006; Mayr, 2013). Hence, if the reader is willing to accommodate the presupposed information, too much (and far-fetched) reasoning needs be made (e.g., extra assumption that every street had 4 policemen). We included this type of bad control items, so as to minimize their differences from the target items, which will make the relevant comparison more direct and reliable. These differences are even smaller compared to those between the relevant test and bad control conditions in experiment 1c. Recall that we ran experiment 1c in order to have the fewest differences possible between test and control conditions. Although the researcher's statement in experiment 1c was made identical across test and control conditions, there were still considerable differences in the corresponding interviewer's questions: there was a different type of question per condition (see example item (7)). This is not the case in the present experiment, where the researcher makes a similar statement and the interviewer posits a similar question across test/bad control conditions.

The experiment also included ten good control items, which were similar to those used in experiment 1a. I am giving an example and its English translation below.

(14) ***Onderzoeker:*** *De dichter des Vaderlands heeft heel mooie gedichten gepubliceerd.*
Researcher: 'The Poet of the Fatherland has published very nice poems.'

Interviewer: *Waar staan die gedichten dan?*
Interviewer: 'Where have the poems been published?'

The interviewer's question above is a neutral and plausible follow-up to the researcher's statement.

Experimental items were fourteen in total (appearing in 7 conditions includ-

ing disjunction items). There were another fourteen different items belonging to experiment 2b, to be described in section 3.4.2 (parts A.4 and A.5 of Appendix A include the experimental items of both experiment 2a and experiment 2b, respectively). Twenty filler items were also included. The total number of stimuli was 58, plus another four trials that were practice items similar to those in experiments 1a, 1b, and 1c.

Experimental items were rotated through seven lists so that each participant saw each item only in one condition. Good controls and fillers were the same across lists.

3.4.1.2 Predictions

First, given the findings of the experiments in the previous section, we expect that the interviewer's statements will give rise to variation effects in the present experiment too. If these effects are derived as underspecified variation effects in superlative items, as Coppock & Brochhagen's (2013b) account and Nouwen's (2015) anti-specificity-based proposal predict, then we should not find a difference in felicity ratings between the *minimum value* and the *higher value* conditions. The underspecified variation reading of the example item, for instance, says that there are at least two distinct numbers ≥ 4 such that there is (at least) one street with that many policemen, without specifying which numbers these are. Thus, on such a variation reading, the interviewer's presupposed information in the *minimum value* items, i.e., *some street has four policemen*, is not justified given the relevant inference conveyed by the researcher. In other words, the variation inference does not license the presupposition *some street has four policemen*. The same holds for the *higher value* items where the interviewer presupposes the non-given and unmotivated information that *some street has five policemen* (regardless of the type of variation inference). Again, here, the underspecified variation inference which requires any two values ≥ 4 to be true of a street with guarding policemen does not entail that there is a street with five policemen. Remember that presenting a presupposition trigger without support for the presupposition leads to lower felicity ratings. Hence, participants should treat *minimum value* and *higher value* conditions similarly in that case, and should accommodate the presupposed information in the relevant interviewer's questions of the two conditions to the same extent.

If, on the other hand, the variation effects conveyed by the researcher's statement belong to the specific variation type, as predicted by the Quantity-based proposals in Büring (2008); Kennedy (2015); Mayr (2013); Nouwen (2015), we expect that the interviewer's question will be judged as more felicitous in the *minimum value* than in the *higher value* items. As said, the specific variation inference further specifies that *some street has four policemen*, which clearly licenses the interviewer's presupposition *some street has four policemen* in the former items. Hence, participants are more likely to accommodate the presupposed information in the *minimum value* items than in the *higher value* items where the interviewer's presupposition *some street has five policemen* does not

follow from the relevant variation inference. This will result in higher felicity ratings for the former than for the latter condition.

Moreover, the presupposed information of the interviewer's question in the *some* control items cannot be made available by calculating an implicature in the researcher's statement, because no scalar inferences are triggered in such embedding contexts. As already said, in order to make the interviewer's questions felicitous, participants would then have to accommodate too far-fetched and unmotivated non-given information. Thus, participants are expected to judge the *some* bad control items as highly infelicitous. We also expect to find a contrast with the *minimum value* (and the *higher value*) items, where the likelihood of accommodating the relevant information to make the question felicitous is higher.

Lastly, given the indications experiments 1a and 1b revealed as to the strength difference of variation effects of superlative and comparative NMs, we expect to find a contrast between these conditions in the present experiment too, contrary to Mayr's (2013) predictions.

Table 3.2 summarizes the predictions of the theoretical accounts in relation to the type of variation effects.

Theoretical accounts	**Predictions**
Underspecified variation Coppock & Brochhagen (2013b), Nouwen (2015) (anti-specificity only)	*minimum value* = *higher value*
Specific variation Büring (2008), Kennedy (2015) Mayr (2013), Nouwen (2015)	*minimum value* > *higher value*

Table 3.2: Predictions of theoretical accounts in relation to the type of variation effects. Symbols > and = stand for higher and comparable felicity ratings, respectively.

3.4.1.3 Participants

Forty-four people filled in the on-line experiment. Only thirty-eight of them (34 female, 1 male, 3 without gender info) were native speakers of Dutch, whose data were kept for further analysis. All participants were undergraduate students at the Faculty of Arts, University of Groningen (most of them first-year students). They received a course credit for their participation and were naive as to the purpose of the study.

3.4.1.4 Procedure

The procedure was as in experiment 1a (and 1b and 1c). Participants had to judge how well the interviewer understood the researcher's claim by giving their response on a -3 to 3 Likert scale, where -3 stands for "the claim is not understood" and 3 for "the claim is understood". Four practice items were provided at the beginning of the questionnaire to familiarize participants with the task, which were as in experiment 1a.

The experiment was created in Ibex and was hosted on Ibex farm.

3.4.1.5 Results

As in experiments 1a-c, the obtained data were analyzed with ordered probit models using the `ordinal` package in `R`. We conducted two analyses: (i) an analysis including the type of NM (superlative vs. comparative) and the type of embedding nominal operator together with the interviewer's question (*minimum vale*, *higher value*, *some bad controls*) as predictors, with the reference levels being the superlative modifier and the *minimum value* conditions, (ii) an analysis where the superlative plus *minimum value* condition was compared to good controls. The analyses also included intercept and slope random effects for participants. We applied backward model selection for random effects and in the following I am reporting on the output of the models with the maximal random-effect structure that converged and had the best fit.

The box plots in figure 3.4 present the obtained felicity ratings per (combination of) condition(s) (see example items in (12), (13), and (14) for the test conditions, the *some* bad controls and for the good controls, respectively). Both the *some* bad controls and the good controls scored as expected, with the former scoring low on the felicity scale (superlative: $mean = -2.197$, $median = -3$, comparative: $mean = -2.026$, $median = -3$), and the latter scoring very high on the scale ($mean = 2.205$, $median = 3$). The findings as to the *some* bad control conditions confirm the observation that scalar implicatures are not available with superlative and comparative modified numerals embedded in the scope of existential nominal quantifiers (Fox & Hackl, 2006; Mayr, 2013). If that were the case, the calculation of these implicatures would license the interviewer's presupposition in those items and readers would find them good. On the contrary, they found them very bad and they did not even make an effort to accommodate the presupposed information in some way or another.

The first analysis (see (i) on previous page) revealed a negative effect of the *higher value* condition ($\beta = -.5529$, $SE = .261$, $z = -2.029$, $p < .05$), such that the superlative items received significantly lower felicity ratings in the *higher value* condition than in the baseline *minimum value* condition. This analysis showed yet another negative effect, that of the *some* bad controls on felicity judgements ($\beta = -2.280$, $SE = .359$, $z = -6.355$, $p < .0001$): superlative items were rated as (highly) significantly less felicitous in the *some* bad control condition than in the *minimum value* condition. Moreover, the effect of

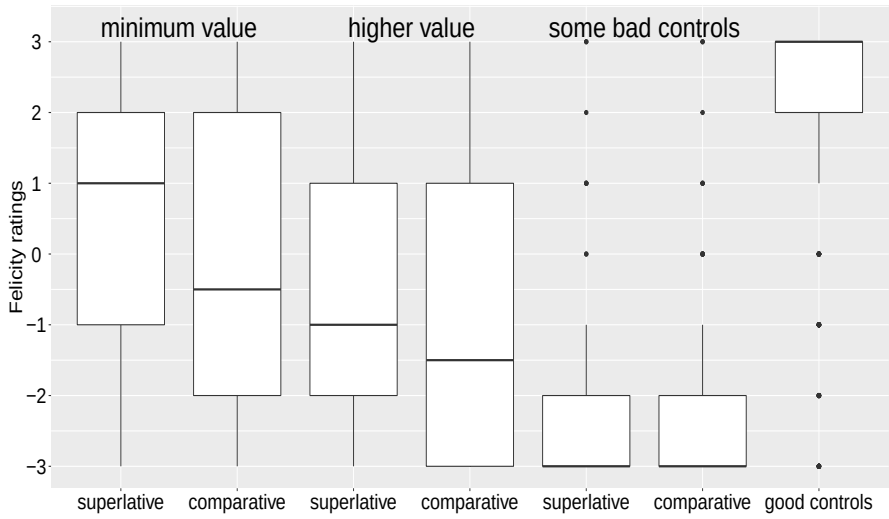

Figure 3.4: Box plots of felicity ratings per combination of conditions in experiment 2a.

NM on felicity ratings, that is the difference between the two NM conditions in the *minimum value* condition (see figure 3.4), did not reach significance ($\beta = -.248$, $SE = .249$, $z = -.994$, $p = .320$). No interaction effect was attested either.

The second analysis (see (ii) on previous page) confirmed that the difference we observe between the baseline condition of the first analysis (superlative + *minimum value*) and the good controls as to the felicity ratings they obtained was highly significant ($\beta = 1.638$, $SE = .267$, $z = 6.148$, $p < .0001$).

Linear mixed-effects regression analysis conducted on the (log-transformed) response times revealed no significant effect whatsoever.

3.4.1.6 Discussion

The difference identified between the *minimum value* and the *higher value* condition, with the former receiving significantly higher felicity ratings, indicates that the interviewer's question was considered to make more sense when it targeted the minimum as opposed to a higher number compatible with the researcher's statement containing the modified numeral phrase. Going back to the relevant predictions, we infer that this finding goes against an account that derives an underspecified variation implication with superlative modified numerals, where any two distinct values (≥ minimum value) are true of an x (Cop-

pock & Brochhagen, 2013b; Nouwen's (2015) anti-specificity-based proposal). This is so, because the interviewer's presupposed information that a higher value is witnessed (*higher value* condition) was not as easily accommodated by participants as her presupposition that the minimum value is witnessed (*minimum value* condition), as would be predicted on an underspecified variation interpretation of the researcher's statement. We can moreover conclude that the finding in question is in favor of an account that derives a specific variation implication with superlative modified numerals, which specifically requires that the minimum value compatible with the researcher's statement is witnessed (Quantity-based accounts in Büring, 2008; Kennedy, 2015; Mayr, 2013; Nouwen, 2015; Schwarz, 2012; Schwarz & Shimoyama, 2011). Such an account actually predicts the attested difference between the *minimum value* and the *higher value* conditions, as the presupposition that the minimum value is witnessed comes directly from a specific variation interpretation of the researcher's statement, while the information that the *higher value* is witnessed has to be accommodated via further reasoning.

Given our above conclusions and that the different types of (partial) variation effects are the result of the different (types of) alternatives that enter the relevant pragmatic mechanism (see sections 2.3.1 and 3.4), we can further conclude that the implicature mechanism responsible for the derivation of the attested specific variation effects engages two alternatives $\forall[n]$ and $\forall[n+1,\ldots)$ for a statement $\forall[n,\ldots)$.

Nevertheless, one might perhaps argue that participants find *minimum value* items as more felicitous than *higher value* items in the baseline superlative NM condition, simply because the relevant interviewer repeats the same number as the researcher in the former but not in the latter items. If this is the case, then we should not find such a difference in the comparative NM condition, since there the interviewer never repeats the number that the researcher states. A post-hoc analysis revealed that the difference between *minimum value* and *higher value* conditions (see also relevant difference in figure 3.4) is significant too ($z = -2.572$, $p = .027$). Therefore, the relevant explanation would be untenable for the attested difference under consideration.

Furthermore, the observed highly strong difference of *minimum value* items from the (minimally differing) *some* bad control items confirms the availability of specific variation implications in the former items. That is, while in the researcher's statements of the latter items the modified numeral does not trigger a Quantity-based reasoning, it does trigger an inference to be derived as a secondary Quantity implicature in the researcher's statements of the former items (which further licences the presupposition in the interviewer's question). Also, the equally strong difference of *minimum value* items from the good controls consisting of neutral researcher-interviewer interactions suggests that specific variation effects are non-obligatory implications, thereby corroborating their pragmatic status.

Importantly, we did not find a significantly different behavior between the two NM conditions, hence based on that we cannot conclude that the specific

variation effects associated with comparative modified numerals are less robust than those of superlative modified numerals (cf. findings of experiment 1c). However, we do observe a tendency in a direction analogous to that found in experiments 1a and 1b: as figure 3.4 shows, superlative items were generally judged as more felicitous than comparative items in the *minimum value* condition. Crucially, this difference was found to be significant in the second half of the experiment, i.e., experiment 2b, testing the same lower-bound superlative modifiers in the scope of modal quantifiers. Experiment 2b sets out to investigate modal variation effects of modified numerals interacting with a universal modal operator. The next section presents experiment 2b, omitting the details that this experiment shares with experiment 2a.

3.4.2 Variation experiment 2b

3.4.2.1 Methods

Experiment 2b uses the researcher-interviewer setup, too, and differs from experiment 2a in that the embedding operator in the researcher's statement is a modal verb. So the corresponding manipulations are: (i) the NM type, i.e., *superlative* (*minstens/ten minste* 'at least') vs. comparative (*meer dan* 'more than'), and (ii) the type of embedding modal operator, i.e., *moeten* 'must' vs. *mogen* 'may, can', together with the interviewer's question.[6] These factors were manipulated in a 2×3 design.[7]

The design and purpose of the conditions of this experiment were the same as in experiment 2a, but completely different items from those of experiment 2a were included. Below I am presenting an example item in all six conditions of interest: (15) illustrates the four test conditions, while (16) (on the next page) constitutes an example of the bad control condition.

(15) **Original item in Dutch**

Onderzoeker: *Een verzekeringsagent moet* $\left\{\begin{matrix}\text{minstens negen}\\ \text{meer dan acht}\end{matrix}\right\}$ *klanten*

researcher an insurance agent must $\left\{\begin{matrix}\text{at least nine}\\ \text{more than eight}\end{matrix}\right\}$ clients

hebben.
have

Interviewer: *Wie bepaalde dat het hebben van* $\left\{\begin{matrix}\text{negen}\\ \text{tien}\end{matrix}\right\}$ *klanten*

interviewer who established that the have.INF of $\left\{\begin{matrix}\text{nine}\\ \text{ten}\end{matrix}\right\}$ clients

genoeg is?

[6] Only one item used *ten minste* (lit. at least) in the superlative condition.

[7] There was also a seventh condition with disjunction, as in experiment 2a, which is not relevant to the purpose of our study.

enough is

Translated item in English

Researcher: An insurance agent has to have $\left\{\begin{matrix}\text{at least nine}\\ \text{more than eight}\end{matrix}\right\}$ clients.

Interviewer: Who established that having $\left\{\begin{matrix}\text{nine}\\ \text{ten}\end{matrix}\right\}$ clients is enough?

The researcher's statement expresses an obligation holding for an insurance agent. The underspecified variation interpretation of the researcher's statement is that there is no specific number m such that in all deontically accessible worlds an insurance agent has m-many clients. The specific variation interpretation additionally requires that there is (at least) one permissible world where an insurance agent has (exactly) nine clients, or in other words, it posits that it is sufficient/OK for an insurance agent to have nine clients. In a similar fashion to experiment 2a, the interviewer's question presupposes that it is OK for an insurance agent to have n clients. Hence, when n is nine (*minimum value* condition), and if the researcher's statement expresses specific variation, the interviewer's presupposition is justified. If the researcher's statement conveys underspecified variation, the relevant presupposition is not justified and should be accommodated. Now, if n is ten (i.e., *higher value* condition), the presupposed information in the interviewer's question is not justified whichever variation inference we derive in the researcher's statement, i.e., specific or underspecified variation. Therefore, participants have to accommodate the non-given information in either case in the *higher value* condition.

(16) exemplifies the bad control condition with the existential modal verb *mogen* ('may'), to be referred to as *may* bad control condition in the following.

(16) **Original item in Dutch**

Onderzoeker: *Een verzekeringsagent mag* $\left\{\begin{matrix}\text{minstens negen}\\ \text{meer dan acht}\end{matrix}\right\}$ *klanten*
researcher an insurance agent may $\left\{\begin{matrix}\text{at least nine}\\ \text{more than eight}\end{matrix}\right\}$ clients
hebben.
have

Interviewer: *Wie bepaalde dat het hebben van meer dan*
interviewer who established that the have.INF of more than
negen klanten teveel is?
nine clients *too much* is

Translated item in English

Researcher: An insurance agent may have $\left\{\begin{matrix}\text{at least nine}\\ \text{more than eight}\end{matrix}\right\}$ clients.

> **Interviewer:** Who established that having more than nine clients is too much?

The interviewer's question in the *may* bad control items presupposes that having more than nine clients is supposed to be a lot for an insurance agent. This presupposition could be licensed if the researcher's statement gave rise to the scalar implicature *it is not permissible to have more than nine clients*, which together with the assertion that it is permissible that the number of clients an insurance agent has is in the range of $[9, \ldots)$ implies that having more than 9 clients is more than is allowed, hence it's too much. However, as already said, modified numerals do not give rise to such implicatures in the scope of existential operators, so *may* bad controls are expected to be judged as highly infelicitous. Clearly, the idea behind these items is the same as in experiment 2a, so the predictions for each condition are also the same (see section 3.4.1.2). Table 3.3 (see next page) repeats table 3.2 from section 3.4.1.2 and presents a summary of the predictions of the theoretical accounts.

Theoretical accounts	**Predictions**
Underspecified variation Coppock & Brochhagen (2013b), Nouwen (2015) (anti-specificity only)	*minimum value* = *higher value*
Specific variation Büring (2008), Kennedy (2015) Mayr (2013), Nouwen (2015)	*minimum value* > *higher value*

Table 3.3: Predictions of theoretical accounts in relation to the type of variation effects. Symbols > and = stand for higher and comparable felicity ratings, respectively.

Similarly to the experimental items with nominal quantifiers, the (different) experimental items of this experiment were fourteen (appearing in seven conditions including disjunctive items). The good controls included in the analyses and the rest were as detailed in section 3.4.1.1.

3.4.2.2 Results & discussion

We carried out the following two analyses here too: (i) one including the NM (superlative vs. comparative) and the type of embedding modal operator together with the interviewer's question (*minimum value*, *higher value*, *may* bad controls) as predictors, with superlative NM and *minimum value* conditions being the reference levels, and (ii) one analysis where we compared these reference levels to the good control items. The analyses were conducted as described

for experiment 2a in section 3.4.1.5.

The box plots in figure 3.5 show the felicity ratings that each (combination of) condition(s) obtained (see (15) and (16) for the corresponding example items).

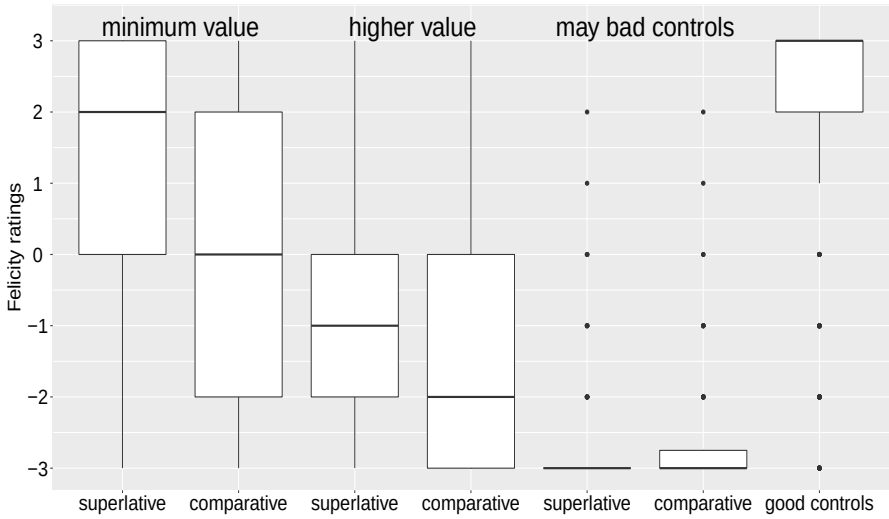

Figure 3.5: Box plots of felicity ratings per combination of conditions in experiment 2b.

In this experiment too, the condition of bad controls (i.e., *may* bad controls) scored as low as expected, showing once again that modified numerals in the scope of an existential operator—as in the researcher's statements of the relevant items—do not trigger a scalar implicature, which would license the presupposition in the interviewer's questions. The good controls are those we already discussed in experiment 2a.

The first analysis (see (i) above) showed a highly significant negative effect of the *higher value* condition ($\beta = -2.524$, $SE = .480$, $z = -5.257$, $p < .0001$), that is, superlative items were judged to be much less felicitous in the *higher value* than in the baseline *minimum* value condition. It was also found that the difference in felicity ratings between the baseline *minimum value* condition and the condition of *may* bad controls was highly significant (negative effect of *may* bad controls: $\beta = -6.325$, $SE = 1.351$, $z = -4.681$, $p < .0001$). Crucially, the first analysis revealed a significant difference between superlative and comparative items in the *minimum value* condition (negative main effect of NM: $\beta = -1.509$, $SE = .470$, $z = -3.208$, $p < .01$) as well as a significant interaction between NM and *higher value* ($\beta = 1.196$, $SE = .521$, $z = 2.298$,

$p < .05$), suggesting that the difference between *minimum value* and *higher value* conditions is bigger for superlative items than for comparative items. Moreover, a post-hoc analysis for the separate NMs showed that there was also a significant difference between the *minimum value* and the *higher value* condition in comparative items (main effect of *higher value* in comparative NM: $\beta = -1.327$, $SE = .379$, $z = -3.505$, $p = .001$).

The second analysis of response data (see (ii) above) corroborated that the difference between the baseline condition of the first analysis (i.e., superlative + *minimum value*) and the condition of good control items, visible in figure 3.5, was significant ($\beta = .766$, $SE = .300$, $z = 2.557$, $p < .05$). Lastly, the analysis conducted on the response time data did not reveal any significant result.

As when in the scope of a universal nominal quantifier, superlative as well as comparative modified numerals are found to trigger specific partial variation effects rather than underspecified variation effects when in the scope of a universal modal quantifier. This follows from the reliable negative main effect of *minimum value* attested (see first and post-hoc analyses). Although the results of experiment 2b come to confirm all significant results of experiments 2a, as already previewed, they further deliver a significant difference between the two NMs and a significant interaction between NM and *higher value*. These findings demonstrate that specific partial variation inferences are more robustly available with superlative modified numerals than with comparative modified numerals. In other words, the inference of the witnessed possibility of the minimum value compatible with the modified numeral, which defines the specific partial (modal) variation, is more robustly associated with superlative than with comparative modified numerals. This fits well with the relevant findings of experiments 1a and 1b, and possibly with the response time data of experiment 1c, too.

Additionally, the difference of the *minimum value* items from the minimally differing *may* bad control items and from the good control items evidence the availability of specific variation inferences and suggest that those inferences are non-obligatory or pragmatic in nature.

As we discussed in experiment 2a as well as at the beginning of this chapter, by experiments 2a and 2b we can further draw conclusions about the pragmatic mechanism numeral modifiers involve. Given the link between input alternative propositions and output inferences, and given our findings, it is concluded that the implicature mechanism responsible for the derivation of specific partial (modal) variation effects of modified numerals like *at least n* and *more than n-1* engages two alternatives $\Box[n]$ and $\Box[n+1,\ldots)$. Perhaps the strength difference in specific variation implications attested between the two modified numerals has to do with the source of these two alternatives and with how tied those are to the meaning of the modified numeral. In any case, finding out why the strength difference in question is not observable when the embedding universal operator is nominal as opposed to modal (note that the relevant experiments had different experimental items) is beyond the scope of this thesis.

Finally, the results concerning the superlative numeral modifier are once

again against Coppock & Brochhagen's (2013b) account and Nouwen's (2015) anti-specificity-based proposal, which both derive underspecified partial variation. They moreover lend support to Quantity-based accounts like those in Büring, 2008; Kennedy, 2015; Mayr, 2013; Nouwen, 2015; Schwarz, 2012; Schwarz & Shimoyama, 2011. However, the attested strength difference between superlative and comparative modifiers challenges Mayr's (2013) prediction that both modifiers trigger variation implicatures to the same extent.

3.5 Overall conclusions

Chapter 3 reported on the first successful attempts to experimentally investigate the likelihood, the strength and the source of the understudied partial variation implications of modified numerals. Taken together, our experiments show that superlative, disjunctive as well as comparative modified numerals give rise to non-obligatory pragmatic variation implications. This confirms Mayr's (2013) prediction that such inferences are available with both superlative and comparative modifiers, and is at odds with the implications of Coppock & Brochhagen's (2013b), Kennedy's (2015), and Nouwen's (2015) accounts that these inferences are not available with comparative modifiers. More importantly, our study further sheds light on the type and source of variation inferences of modified numerals, revealing that they necessarily include an inference about the minimum value compatible with the modified numeral (*specific variation*). This goes against Coppock & Brochhagen's (2013b) account and Nouwen's (2015) anti-specificity-based proposal that derive underspecified partial variation. Interestingly, the results of experiment 1a, 1b, and 2b demonstrate a less robust pragmatic nature for the partial variation effects associated with comparative modifiers. Although the felicity ratings obtained in experiments 1c and 2a did not corroborate this finding, the response time data of experiment 1c as well as the numerical difference between the two NM conditions attested in experiment 2a could be taken to point to such an explanation. Given all this, I am inclined to believe that there is a difference between comparative and superlative modified numerals as to the strength of the specific variation inferences they trigger. This is, furthermore, contrary to Mayr's (2013) prediction that these implicatures arise with both types of numeral modifiers to the same extent.

The present chapter, therefore, comes to dispel a prevalent misconception in the literature that superlative modifiers, or class B modifiers in general, trigger inferences that comparative modifiers do not. Our findings do confirm that the two classes of numeral modifiers are different, but only in the strength, not in the type of inferences triggered. Given that, one could say that Nouwen's (2010) class A/B distinction or as this was reformulated in Nouwen (2015) accounting for inferences that are available *only* with class B numeral modifiers is misguided to some extent. Although this can now be argued regarding variation inferences only, in chapter 6 a similar conclusion will be made as far as

ignorance inferences are concerned.

Crucially, that specific variation is the right type of partial variation that modified numerals trigger gives significant insight into the pragmatic mechanism behind this inference and the type of alternatives that are fed into this mechanism. In particular, it demonstrates that for an utterance with an embedded modified numeral in the scope of a universal operator, $\forall/\Box[n,\ldots)$, the two relevant alternatives that are being negated in the pragmatic reasoning (exhaustification output) are $\forall/\Box[n]$ and $\forall/\Box[n+1,\ldots)$. This is the case for a Quantity-based account à la Büring (2008) and Kennedy (2015), which posit disjunctive alternatives for superlative modifiers, or à la Mayr (2013), Nouwen (2015), or Schwarz (2012), which posit a more articulate set of alternatives. While in the former type of accounts the alternatives are generated by substituting only the modifier with elements from < *exactly*, *more than* >, in the latter accounts the alternatives are generated by substituting both the numeral and the modifier (see **Specification of alternatives** in section 2.2.2 for the specifics of each account). Given that comparatives trigger specific variation too, this further implies that a similar set of alternatives should be assumed for a comparative modified numeral like *more than n* as well, i.e., either {exactly $n+1$, more than $n+1$} or a more complex one, as in Mayr (2013) (see (17)) or as in Ciardelli et al. (2017) (see (18)).

(17) {more than $n+1$, more than $n+2$, ..., fewer than $n+1$, fewer than $n+2$, ...}

(18) {exactly $n+1$, exactly $n+2$, ..., more than $n+1$, more than $n+2$, ...}

Given that superlative and comparative modifiers are associated with a similar set of alternatives, what creates the resulting strength difference in variation effects? It could have to do with the source of these alternatives, as I have already hinted at earlier. That is, it might have to do with how tied to the meaning of the modifier the relevant alternatives are. This possibility will further be tackled in chapter 6, also in relation to the overall experimental findings up until that point and data like those discussed in section 2.2.5 of chapter 2. Also, the strength difference in question will have a key role in the setup of chapter 6 and the experiment therein. Importantly, the specific variation effects that we found evidence for in the present chapter (as opposed to underspecified variation) will play a crucial role in the experiments to come (chapters 5 and 6), but then in the form of specific speaker ignorance effects, the underlying idea being that the same pragmatic mechanism lies behind the derivation of these two types of inferences.

This chapter has answered the first bunch of questions raised in section 1.4 of Introduction (chapter 1), repeated below, by showing that variation effects are non-obligatory pragmatic inferences that become available with class B (superlative and disjunctive) modifiers as well as with comparative modifiers, only they tend to be less robust in the latter case. More specifically, they are shown to be of the specific variation type, thereby indicating that $[n]$ and $[n+1,\ldots)$

(where n is the minimum value compatible with the modified numeral) are the alternatives fed into the relevant pragmatic mechanism. This is a good start laying important foundations for delineating the pragmatics of modified numerals.

- ▷ What is the likelihood of drawing a variation inference? Do different (truth-conditionally equivalent) modified numerals give rise to variation effects to the same extent? What is the semantic/pragmatic status and strength of variation effects? What type of variation do modified numerals involve and what are the implications for their underlying mechanism of derivation?

In the following, I turn to the most well-studied type of inference associated with modified numerals, i.e., speaker ignorance. In chapter 4, I present the existing experimental literature on ignorance inferences, while in chapters 5 and 6 I detail my own contribution to that literature.

CHAPTER 4

Experimental background to speaker ignorance effects

4.1 Introduction

A quite large series of experiments has been carried out to investigate the extent to which numeral modifiers give rise to ignorance inferences and whether these arise semantically or via pragmatic reasoning.[1] Exploring ignorance effects by means of experiments could further help figure out how, or mostly whether, superlative modifiers are different from comparative modifiers in that respect.

The various experimental investigations differ in whether they set out to test speaker ignorance inferences directly or indirectly, as well as in the designs and methodologies used. The vast majority of the experiments employ off-line methods, probing ignorance effects indirectly by comparing statements with modified numerals to some given definite value, provided by a variety of ways: (i) as the premise in a reasoning task (Geurts & Nouwen, 2007; Geurts, Katsos, Cummins, Moons & Noordman, 2010), (ii) as a picture in a truth-value judgment task (Coppock & Brochhagen, 2013a), or (iii) as part of a short discourse (Cummins & Katsos, 2010; McNabb & Penka, 2015). Contrasts in responses between superlative and comparative conditions are then attributed to the speaker ignorance effects of superlative modifiers. There has also been some attempt to record and look into participants' reading and response times in a judgement task like those just mentioned (Cummins & Katsos, 2010), hypothesizing that the availability of ignorance effects adds to the core meaning of the modifier creating some complexity, which is expected to be reflected in the

[1] This chapter constitutes an expansion of section 2 of Nouwen, Alexandropoulou & McNabb's (2018) chapter in the Handbook of Experimental Semantics and Pragmatics.

time measures. Moreover, there are a few recent studies with more direct experimental investigations of speaker ignorance effects, which incorporate epistemic information in the experimental setup (McNabb & Penka, 2015) by manipulating the speaker's epistemic state in some of their experimental conditions, or that manipulate the question under discussion (QUD) in an experiment that combines an off-line and an on-line task (Westera & Brasoveanu, 2014).

The findings of the experimental research under consideration provide evidence that ignorance inferences have a pragmatic status rather than a semantic one. However, this evidence is mostly indirect. What's more, the little direct evidence that does exist turns out to be rather controversial. Given the existing experimental findings, it will be concluded that more detailed direct study is needed, which would ideally look into the incremental interpretation of numeral modifiers. Such an investigation is argued to yield more firm indications and conclusions about the nature of speaker ignorance inferences, contributing more enlightening information about what kind of accounts of numeral modifiers best capture those inferences.

In the following, I will go through all the aforementioned studies in some more detail. I will first discuss the studies indirectly testing for speaker ignorance effects (section 4.2.1) and then turn to more direct investigations of such effects (section 4.2.2). Finally, I will conclude this chapter by reflecting on those studies and the findings thereof, and by highlighting what is still missing from the experimental investigation of speaker ignorance effects of numeral modifiers (section 4.3), paving the way to my first experiment on speaker ignorance (chapter 5).

4.2 Empirical research on speaker ignorance effects of numeral modifiers

4.2.1 Indirect investigation of ignorance inferences

Starting with the studies testing indirectly for ignorance, I should first mention Geurts & Nouwen's (2007) experimental findings. The off-line investigation they report on to supplement and back up their theoretical proposal seems to be the first experimental study where ignorance effects indirectly prove to be available with superlative numeral modifiers, but not with comparative modifiers. Specifically, Geurts & Nouwen carried out two paper-and-pencil experiments (in Dutch) testing the inference patterns of the relevant superlative and comparative numeral modifiers. Participants had to judge the validity of arguments like the following (premise in (1) and conclusion in (2)):

(1) Beryl had n sherries.

(2) Beryl had { *at least* n / *at most* n / *more than* $n-1$ / *fewer than* $n+1$ } sherries.

Conclusions with comparative modifiers were almost unanimously accepted, while conclusions with superlatives were judged as valid conclusions given the premise in 22% of the cases for *at most* and in half of the cases for *at least*. These low percentages of the superlative conditions and their difference from the respective comparative conditions can be attributed to the fact that superlative modifiers give rise to ignorance and thus the relevant conclusions are at odds with the competence and certainty conveyed by a statement with a definite value as in (1). The authors, under the assumption that this method tests for entailment relations, interpret this result in favor of the semantic account they propose, whereby ignorance is encoded in the lexical semantics of superlative modifiers.

Geurts et al. (2010) in a very similar experiment (also conducted in Dutch) found a similar contrast between comparative and superlative conditions: Although the acceptability of arguments with comparative numeral modifiers was at ceiling, only half of the participants judged the arguments with the corresponding superlative modifiers as valid. Again, these findings could be attributed to the availability of ignorance inferences with superlative modifiers that clash with the knowledgeability signaled by (1). More specifically, Geurts et al. take the observed contrast to indicate a semantic difference between comparative and superlative modifiers in favor of the semantic account put forth by Geurts & Nouwen (2007). Furthermore, this contrast between superlative and comparative modifiers manifests itself also in an on-line verification study reported in Geurts et al. (2010). This is the first attempt ever to look into the on-line judgement of sentences with a numeral modifier in a verification task with an on-line design. More accurately, Geurts et al. measured how much time participants spend reading a statement with a numeral modifier as in (3) (reading times), as well as how long it takes them to decide whether that statement is true or false of a situation that was later revealed (response times), where up to four identical X's appeared, see (4).

(3) There are $\left\{\begin{array}{c}\text{more than 2}\\ \text{at least 3}\\ \text{fewer than 3}\\ \text{at most 2}\end{array}\right\}$ X's. (where X is the letter *A* or *B*)

(4) X / XX / XXX / X X X X

They found no effect in the reading times, but there was a reliable difference between comparative and superlative conditions in decision times, with the latter delaying (true/false) judgements. The authors do not report any results on the percentage of true/false responses or the response times of true and false judgements separately. Nonetheless, they take their finding to indicate that superlative modifiers are harder to process compared to comparative modifiers, supporting Geurts & Nouwen's (2007) account, where superlative modifiers are semantically more complex because they additionally incorporate an epistemic component.

Both Geurts & Nouwen (2007) and Geurts et al. (2010) assume that their off-line method exclusively tests for entailment relations, thus they interpret their results in favor of the semantic account offered by Geurts & Nouwen (2007). However, intuitively it seems to be the case that their methodology might test for any type of implication, either semantic or pragmatic. As Coppock & Brochhagen (2013a) specifically argue, the truth-conditional content is not the only type of information that might determine judgements in a setting where what matters is whether the conclusion and the premise(s) share (any kind of) common information, be it truth-conditional or pragmatic (see also Ariel, 2004 on the difference between lexical semantic context and semantic-pragmatic truth compatibility). Therefore, it is very likely that pragmatics might have interfered with Geurts & Nouwen's (2007) and Geurts et al.'s (2010) inference-validity judgement tasks, calling into question the authors' interpretation of the relevant findings in favor of Geurts & Nouwen's (2007) semantic account of speaker ignorance. Furthermore, it is not clear what the longer decision times incurred by the superlative conditions in Geurts et al.'s (2010) on-line verification judgement task are due to and what they are to be associated with, especially given that very little information is reported with respect to the results in question. Even less can surely be concluded given the null results reported as regards the more direct measure of reading times in Geurts et al.'s (2010) second experiment.

Coppock & Brochhagen (2013a) set out to directly test the hypothesis that ignorance effects are semantic inferences. To that end, they use a truth value judgement task similar to that used in the poorly reported on-line experiment by Geurts et al. (2010). They argue that this is an unambiguous and better method for identifying entailments compared to the validity tasks, as it blocks pragmatic intrusion. In their task, native speakers of English had to judge whether a sentence like (5) was true or false given a picture of n-many objects of type X.

(5) There are { *at least* n / *at most* n / *more than* $n-1$ / *fewer than* $n+1$ } Xs in the picture.

Although their target sentences were very similar to those used in Geurts & Nouwen's (2007) and Geurts et al.'s (2010) off-line tasks and had to be assessed against a similar context, albeit this time visually depicted rather than written, this task revealed very different results (remember that we do not have any information about the relevant percentages in Geurts et al.'s (2010) on-line verification experiment). All (superlative and comparative) test conditions obtained surprisingly high percentages of *true* judgements, without being different from each other, but crucially being significantly different from the control condition of logical contradictions (i.e., *at least* $n+1$, *at most* $n-1$, *more than* n, *fewer than* n), which obtained very low percentages of *true* judgements, as expected. These results eliminate the possibility that speaker ignorance is part of the lexical semantics of superlative modifiers. Under such an assump-

tion test sentences with superlative modifiers in this task are predicted to be judged as contradictions against a depiction of an exact number of objects and thus pattern with logical contradictions and differently from the corresponding comparative conditions, which are claimed not to trigger ignorance implications. Coppock & Brochhagen (2013a) conclude that their findings go against a semantic account of speaker ignorance effects and are in line with a pragmatic account of them.

In the foregoing, I first presented those experimental studies whose findings witness to the availability of speaker ignorance effects with superlative modifiers as opposed to comparative modifiers and can be accommodated both by a semantic and a pragmatic account of such effects. Coppock & Brochhagen's (2013a) findings come to eliminate the former possibility and are rather compatible with a pragmatic account of speaker ignorance. In the remaining of this section, I will discuss studies that make use of a more elaborate methodology, which enables testing for inferences with a pragmatic status.

In particular, in order to determine whether ignorance inferences are semantic or pragmatic in nature, Cummins & Katsos (2010) employ a method that has been shown to distinguish between logical contradictions and pragmatic infelicities (recall the discussion in section 3.2 of previous chapter). Katsos & Bishop (2011) show that a gradient, as opposed to a binary (as in the previous experiments), response paradigm leads participants to distinguish among entailments, logical contradictions and pragmatic infelicities. In the light of this finding, Cummins & Katsos (2010) use a Likert scale from -5 ("incoherent") to +5 ("coherent"). Their participants, who were native speakers of English, are asked to judge on that scale how coherent an utterance consisting of a pair of sentences is. Target items comprised sentences with a numeral modified by *exactly* as continuations of a statement with a lower- or upper-bound numeral modifier, as in (6).

(6) Jean has $\left\{\begin{matrix}\text{at least / more than}\\ \text{at most / fewer than}\end{matrix}\right\}$ n houses. (target item)

$\left\{\begin{matrix}\text{Specifically}\\ \text{In fact}\end{matrix}\right\}$, she has exactly $\left\{\begin{matrix}n+1 \ / \ n\\ n-1 \ / \ n\end{matrix}\right\}$ houses.

(7) Jean has some houses. (control item)

$\left\{\begin{matrix}\text{Specifically}\\ \text{In fact}\end{matrix}\right\}$, she has $\left\{\begin{matrix}\text{none}\\ \text{half}\\ \text{all}\end{matrix}\right\}$ of the houses.

Cummins and Katsos included both a condition with *specifically* and one with *in fact* (see continuation sentence in (6)), as well as three types of control items with *some*, see (7): logical contradictions (with *none of* in the continuation sentence), logical entailments (with *half of* in the continuation sentence), and pragmatic infelicities (with *all of* in the continuation sentence). They predicted

that logical contradictions will be judged as completely incoherent, scoring at the lower part of the Likert scale, logical entailments would be judged as coherent, with scores at the upper part of the scale, and pragmatic infelicities (or cancellations) would be judged more coherent than the contradictions, but yet less coherent than the entailments, receiving scores from the middle of the scale. Indeed, all control conditions obtained scores as predicted, justifying the effectiveness of the particular method.

Test items, i.e., items with a superlative numeral modifier (*at least* / *at most*) and *exactly* $n+1/n-1$ or *exactly* n in the continuation sentence, received scores higher than the superlative logical contradictions (i.e., *at least* n... *exactly* $n-1$ and *at most* n... *exactly* $n+1$) and lower than their comparative counterparts (i.e., *more than* n... $n+1$ and *fewer than*... $n-1$), which scored as high as the control condition of logical entailments. This difference between superlative conditions and comparative conditions was observed regardless of whether the continuation sentence had *specifically* or *in fact*, with the latter conditions obtaining higher coherence rates in general. The coherence rates the superlative test conditions received (compared, for instance, to the control condition of entailments) can be attributed to the speaker ignorance effects triggered by the first sentence in those conditions, which are in conflict with the knowledgeability conveyed by the relevant continuation sentence with *exactly*. In contrast, the high coherence rates of the comparative conditions suggest that there is no such clash or inconsistency with the information of the respective continuation sentences, hence, no indication of speaker ignorance effects.

Next, the significant difference between the superlative conditions and the superlative logical contradictions indicates that participants find the former more coherent and that the ignorance triggered in those conditions is not a semantic inference that is contradicted by the information of the continuation sentence. Rather, Cummins & Katsos take this difference as evidence that the ignorance implication is a pragmatic inference that is cancelled by the continuation sentence. This inference is further argued to be a less established or robust pragmatic inference compared to the scalar implicature in the control pragmatic condition with *some* in (7), because the cancelled ignorance inferences (superlative test items) were judged as more acceptable than the cancelled scalar implicatures (control pragmatic items). The findings revealed by Cummins & Katsos's (2010) novel methodology allow them to reject a semantic account of speaker ignorance and adequately argue in favor of a pragmatic account.

McNabb & Penka (2015) used Cummins & Katsos's (2010) methodology as well, and they administered a similar coherence judgement task with a (–5 to +5) Likert scale. The only difference was that their target sentences with superlative modifiers had to be judged in terms of coherence against a rich preceding discourse containing an exact quantity, as opposed to the continuation sentence Cummins & Katsos's (2010) paradigm included. (9) presents an example of a target sentence and (8) shows the context the target sentence had

to be judged against.

(8) After the soccer practice, Cassandra sent Eduard to collect all of the soccer balls lying around in the court. He collected the n balls that were scattered around the court. When he was done, he went back to Cassandra and said:

(9) "I collected {*at least* / *at most*} {n-1 / n / n+1} balls."

Similarly to Cummins & Katsos (2010), McNabb & Penka found that utterances with *at least* n and *at most* n were judged to be more coherent than the respective logical contradictions (*at least* $n + 1$ / *at most* $n - 1$) and so were the utterances with *at least* $n - 1$ and *at most* $n + 1$. Once again, the difference from logical contradictions is taken to indicate that the ignorance signaled by superlative modifiers is not an entailment, but is derived via a pragmatic mechanism.[2]

One could also argue that participants judged the superlative conditions as incoherent in that particular task, because they might have found it (prima facie) weird or unreasonable that the speaker, who collected the n-many balls herself given (8)-(9), would not mention the exact n, but rather use a superlative modified numeral. Perhaps participants further tried to accommodate the use of a superlative modified numeral by thinking of possible reasons that made the (knowledgeable) speaker make such an utterance. In other words, it might be the case that the observed difference does not reflect speaker (and agent) ignorance about the exact quantity of balls collected. A plausible interpretation of the observed results could be that the speaker and agent of the ball-collecting activity used a superlative modified numeral to signal that it is not relevant to mention the exact number of balls she collected (which she remembers) or that she does not want to reveal all the relevant information. Note that this experiment did not test the comparative counterparts *more than* and *fewer than*. Under this alternative explanation, the comparative conditions would be predicted to receive as low coherence ratings as superlative modifiers did, the idea being that another message (than speaker ignorance) is being communicated in that case too by not mentioning the (known) exact n.

[2]It is worth mentioning here that McNabb & Penka (2015), and actually Coppock & Brochhagen (2013a) too, found a contrast between *at least* and *at most* conditions, which cannot have to do with ignorance: items with *at most n + 1* were considered significantly less coherent than the similar *at most n*, *at least n*, and *at least n 1* test conditions. The authors explain this asymmetry in the following terms: Participants have difficulties accepting the speaker's utterance *at most n + 1* given a quantity n introduced in the context, because she is stating the possibility that a greater quantity (i.e., *n + 1*) is true while she knows that it is false given n in the preceding context. This is not the case for *at least n*, although this too allows for the possibility of a higher quantity than n that is not true given n in the context. The contrast arises because *at most n+1*, as opposed to *at least n*, explicitly states, or highlights in Coppock and Brochhagen's inquisitive semantic terminology, the possibility that a greater number than n is true.

4.2.2 Attempts of direct investigation of ignorance inferences

As already said, the studies previously presented test indirectly for speaker ignorance inferences. Namely, they do not include an explicit manipulation of speaker ignorance nor do they measure speaker ignorance directly, but rather via measuring something else, e.g., coherence of discourses containing utterances with a potential ignorance-trigger. Only very recently have there been the first few attempts of using methodologies and designs that directly investigate ignorance effects of numeral modifiers and their semantic/pragmatic status. McNabb & Penka (2015) set out to do so in their next experiment, looking into how people interpret superlative modifiers within an ignorance context by manipulating the speaker's epistemic state. Importantly, in this experiment superlative modifiers are embedded under universal and existential deontic modals, as the authors' primary aim was to find out which combinations of superlative modifiers and modals obviate ignorance inferences. The paradigm they used is very similar to that in their previous experiment, including the same Likert scale. This time participants had to judge whether the speaker's utterance (in quotes in (10)) was compatible with the preceding context that sets up her epistemic state, as illustrated in (10-a) and (10-b) below.

(10) a. +knowledgeable speaker
The secretary, who was involved in the selection process, said:

b. –knowledgeable speaker
The secretary apologized for not knowing the requirements for the application, and said:

"You are { allowed / required } to have { *at least* / *at most* } 3 works in the portfolio you send us."

The combinations *allowed to + at most* and *required to + at least* received significantly higher coherence ratings than the combinations *allowed to + at least* and *required to + at most* in the +knowledgeable speaker condition. The authors interpret this difference as showing that the latter combinations trigger ignorance inferences. That is, the +knowledgeability speaker condition clashes with speaker ignorance, so the attested lower coherence rates of the combinations *allowed to + at least* and *required to + at most* in that speaker knowledgeability condition are taken to be an indication of such a clash. These combinations are also found to be significantly more coherent compared to the control condition of logical contradictions, suggesting that the ignorance effects triggered are pragmatic inferences. Note that the control condition of contradictions consisted of completely different items, that is, they differed from the test items in various aspects. As other factors could then be responsible for the attested contrast of the relevant conditions as well, one could argue that this contrast

might not be that telling as to the pragmatic status of speaker ignorance or that we should be more cautious about treating the relevant comparison as a pragmatic diagnostic. Despite that, we could still take the observed difference to indirectly suggest that ignorance is not to be accounted for by a semantic analysis, but rather by a pragmatic one.

Let's now turn to the results of the more important condition of –speaker knowledgeability (see (10-b)), which was included in order to test for speaker ignorance effects directly. The authors predicted that there will be no contrast among the various modal + superlative modifier combinations in this speaker condition and that all of them will receive scores at ceiling, assuming that all combinations have an underlying scope configuration that gives rise to speaker ignorance effects. Alternatively, one could even hypothesize that this condition would yield the mirror contrasts of those revealed in the +knowledgeability condition between the relevant modal + superlative modifier combinations, under the assumption that if a combination is more likely to be compatible with a +knowledgeability context compared to another combination, then the former could possibly be judged as being less compatible with a –knowledgeability context than the latter. Surprisingly enough, the results in this condition turned out to be same as in the +knowledgeability condition: *required to + at most* and *allowed to + at least* still obtained lower coherence rates than the remaining combinations. The authors speculate that this could be because participants possibly made the assumption that the speaker actually has knowledge of the exact requirements or of what is permitted being, for instance, a secretary in the relevant company in (10). As a result participants treated the –knowledgeability condition as a +knowledgeability condition to some extent. Thus, although the –knowledgeability condition was set up as the condition that directly tests for speaker ignorance inferences, the aforementioned confound does not make this possible yet. Another possibility could be that the combinations *required to + at most* and *allowed to + at least* are just more degraded, independently of the condition, and as a result they receive lower scores overall compared to *allowed to + at most* and *required to + at least*. In conclusion, McNabb & Penka's (2015) last study too provides just indirect evidence that speaker ignorance effects are pragmatic inferences, as the condition set up to directly probe ignorance turned out not to be successful.

Westera & Brasoveanu (2014) is another study that directly investigates speaker ignorance, by including an explicit measure of ignorance and an online design. Specifically, Westera & Brasoveanu are the first to carry out an on-line study on speaker ignorance effects of numeral modifiers. They carried out two self-paced reading experiments in English combined with a validity judgement task, which aimed to directly probe ignorance effects with upper-bound numeral modifiers and their context-sensitivity. They make and test the following assumptions. First, in precise contexts, i.e., when there is a *how many* question under discussion (QUD), superlative and comparative modifiers are hypothesized to trigger speaker ignorance effects to the same extent. To elaborate, the idea is that a *how many* question asks for a precise answer and an

interlocutor that uses a superlative or a comparative modifier in her answer signals that she is most possibly unable to give such an answer because of lack of knowledge, as in example (11) below, repeating (2) from chapter 2.

(11) A: How many souvlakis did Magda eat?
B: She ate at least/more than 3 souvlakis.

Next, they assume that if there is an underspecified QUD, as, for instance, in the case, of an out-of-the-blue utterance with a numeral modifier, superlative modifiers are more likely to give rise to speaker ignorance than comparative modifiers. Westera & Brasoveanu base this assumption on corpus findings. Those findings show that superlative modifiers occur more frequently than comparatives in precise contexts, like that in (11). Hence, when people are exposed to an utterance with a numeral modifier in an underspecified context, they are more likely to accommodate a precise context in the case of superlative modifiers, which is the right setting for triggering speaker ignorance. This results in a higher likelihood of an ignorance interpretation for superlative as opposed to comparative modifiers. Lastly, in imprecise or coarse contexts, where an imprecise/rough answer is asked for or is at stake, no ignorance is expected to arise with superlative or with comparative modifiers. The idea is that the range of values expressed by the basic meaning of either modifier satisfies the imprecise QUD.

Participants in Westera & Brasoveanu (2014) had to read short dialogues between a judge and a witness in a court setting, as in (12)–(13). They were first presented with a question posed by the judge, which was displayed as a whole. This question contained the QUD manipulation, see (12-a)–(12-c). Then participants had to read the witness's answer, which contained the superlative modifier *at most* or the comparative *less than* and a main verb of perception, see (13). Crucially, the witness's answer was presented word-by-word and reading times of each word were recorded. Lastly, the witness's statement was followed by the judge's conclusion that the witness did not know the exact quantity of the objects under discussion (see (14)), which evidently amounts to speaker ignorance. Participants had to assess how justified the judge's conclusion was on a Likert scale from 1 to 5, where 1 stands for "not justified at all" and 5 for "strongly justified".

(12) The judge asks:
a. Precise QUD
"How many of the diamonds did you find under the bed?"
b. Underspecified QUD
"What did you find under the bed?"
c. Imprecise QUD
"Did you find at most / less than ten of the diamonds under the bed?"

(13) Witness: "I found { *at most* / *less than* } ten of the diamonds under the bed."

(14) Judge's conclusion: "The witness doesn't know exactly how many of the diamonds she found under the bed."

In the first experiment, participants found the judge's conclusion to be more justifiable when the QUD was precise (*how many* question in (12-a)) or underspecified (*what* question in (12-b)) than when it was imprecise (*polar* question in (12-c)). That is, in the former conditions the witness was judged to be more ignorant than in the latter condition. It was further found that, when just considering the responses to a precise QUD, the judge's conclusion was more justifiable when the witness had used *at most* vs. *less than*. However, when all data were taken together, there was no interaction between the QUD and the numeral modifier factors nor a main effect of the numeral modifier. These findings indicate that the QUD plays a role in how people interpret utterances with numeral modifiers and in the likelihood of speaker ignorance to arise in particular, independently of the type of the numeral modifier. This constitutes the first piece of evidence that comparative numeral modifiers too can trigger speaker ignorance, if we consider the right context. In general, these new findings go against a semantic account of speaker ignorance that would predict no context sensitivity, but also against those accounts that only derive speaker ignorance with superlative modifiers. Note that we have to be cautious with the interpretation of the QUD main effect mentioned at the beginning of this paragraph. The lower ratings the imprecise QUD condition received compared to the other two QUD conditions could as well be due to the unnatural echoic responses to the relevant questions. That is, the witness's utterance in (13) as an answer to the question in (12-c) repeats unnecessary lexical material while the expected answer is either *yes* or *no*. This might have made participants to rate this condition low.

The difference found between *at most* and *less than* for the *how many* questions is not expected given the authors' predictions. Although this is hypothesized to be a precise QUD condition, according to the authors this result implies that the relevant QUD might be underspecified to some extent (recall their relevant assumption based on corpus data). And while the *how many* condition might have been treated as an underspecified QUD condition, the authors further provide corpus data suggesting that *what* questions might have been interpreted more precisely than predicted, making this condition the actual precise QUD condition.

Possibly due to the aforementioned obscurities, in their second similar experiment Westera & Brasoveanu modified somewhat the QUD conditions. (15) illustrates the new conditions, which seem to be more clear-cut and to better and more accurately test for QUD sensitivity.

(15) The judge asks:

a. Imprecise QUD
"Approximately how many of the diamonds did you find under the bed?"

b. Precise QUD
"Exactly how many of the diamonds did you find under the bed?"

c. Precise QUD
"Did you find eight, nine, ten, or eleven of the diamonds under the bed?"

(16) Witness: "I found { *at most* / *less than* } ten of the diamonds under the bed."

(17) Judge's conclusion: "The witness doesn't know exactly how many of the diamonds she found under the bed."

The results showed that, as in the first experiment, the judge's conclusion was considered more justifiable in the precise QUD conditions (*exactly how many* and multiple-choice questions) than in the imprecise QUD condition (*approximately how many* question). In addition, no difference was found between numeral modifiers and no interaction was attested between the QUD and the numeral modifier factors either. In fact, these latter null results are in line with the authors' predictions, as this experiment includes only (well-designed) precise and imprecise contexts, wherein the two types of numeral modifiers are hypothesized to exhibit a uniform performance (remember that an interaction effect was only expected to arise in underspecified contexts, which are not included in the design of the present experiment). As to the main finding of this experiment, i.e., the significant effect of QUD, this again indicates that speaker ignorance inferences are triggered more robustly in some (certain) types of contexts than in others, regardless of the type of the numeral modifier. This context-sensitivity effect suggests that speaker ignorance inferences are not semantic inferences expected to arise across the board.

Let us now turn to Westera & Brasoveanu's on-line tasks. Their first self-paced reading experiment, with the polar, *how many*, and *what* QUD conditions, revealed the following: (i) the words *ten*, *of*, *the*, and *under*—which are self-paced revealed immediately after the numeral modifier—were read slower in the witness's answer to the *what* question in (12-b) compared to the polar question in (12-c), (ii) the same slowdown was found at the words *ten* and *of* in the *how many* condition as compared to the baseline polar question condition, (iii) the word *under* was read slower when a *how many* QUD was answered by the witness using a superlative modifier, but this was not the case for the precise QUD & comparative condition. The authors take these results to confirm the off-line data of the first experiment, linking the increase in reading times in the *what* and *how many* questions vs. the polar questions to the increase of speaker ignorance readings in these conditions in the off-line task. They also link the longer reading times observed in the *how many* & superlative condition to the

difference they find between numeral modifiers in the *how many* condition in the off-line task. Westera & Brasoveanu conclude that the increased reading times observed in the words following the numeral modifier region might be due to the costly on-line calculation of the pragmatic ignorance inference triggered by the numeral modifier or to the relevant silent intonational effects during reading depending on the context. It is particularly striking and remarkable that, in the second on-line task, which included more clear-cut QUD conditions, there was no significant effect whatsoever. The authors attribute this to a possible habituation effect, as in the second experiment their fillers included a judge inferring ignorance on the part of the witness, similarly to the test items.

The findings of this study are taken to show that speaker ignorance inferences are available with both superlative and comparative modifiers, depending on the context, and are context sensitive, hence they cannot be semantic inferences, but they deserve a pragmatic analysis. However, there is a number of concerns that this study raises or does not resolve. First and foremost, it is by no means clear how exactly participants interpreted the target sentences during reading; for instance, whether they derived speaker ignorance inferences on-line in the target sentence or whether they derived speaker ignorance when they were prompted to do so by the judge's conclusion (*The witness doesn't know exactly how many...*). Also, the reading time effects were not replicated in the follow-up experiment of (at least) as much power (Exp1/2: 36 test items, Exp1: 35 participants, Exp2: 51 participants), although the off-line effects were. This fact suggests that the effects in the first on-line experiment should not be linked to the ratings of ignorance inferences, if we ignore the habituation effect assumption, but should perhaps be associated with the particular conditions. Speculating a bit, the use of the definite phrase *the diamonds*, for instance, in the target sentence might sound weird when this constitutes an answer to the *what* question in (12-b), because the relevant referent is not established in the context in that case compared to the other two types of questions in (12-a) and (12-c). The extensive slowdown attested in this condition then might have to do with participants' need to accommodate the relevant referent (note that this slowdown extends to more words compared to the effect in the *how many* condition).

Furthermore, the use of perception verbs (*find, see, hear*) by the witness in the target sentences strongly implies competence on the part of the speaker, which possibly clashes with the use of a numeral modifier in the *what* and *how many* conditions and perhaps even affects reading times. Or it is also possible that given the speaker competence signaled by the use of such verbs participants might have interpreted the use of a superlative modified numeral by the witness in the *what* and *how many* conditions as signaling other reasons of not mentioning the exact number than speaker ignorance (note that witness's ignorance is only imposed after participants have read the target sentence). Such reasons could be that the witness does not think it is relevant to give more/precise information she possesses or that the witness does not want to do so. Although the authors explicitly state that they chose the court setting

to ensure that *the witness has nothing to hide and is fully cooperative* (Westera & Brasoveanu, 2014, p. 421), this does not mean that they were successful in doing so; cf. a familiar example in the pragmatic literature in (18), containing a dialogue in a court setting between Prosecutor (P) and Broston (B), where the latter gives no or partial information and avoids saying that he had a Swiss bank account.

(18) P: Do you have any bank accounts in Swiss banks, Mr. Bronston?
B: No, sir.
P: Have you ever?
B: The company had an account there for about six months, in Zurich.
(adapted from Asher & Lascarides, 2013)

Even if we assume that the choice of the court setting was effective in Westera & Brasoveanu's (2014) reading task, this still does not exclude the possibility that the witness gives partial information (by using a modified numeral vs. a bare numeral) because she does not think it is relevant to say the exact number, while she believes that the maximum value/threshold she conveys is already informative enough; note also that all test items contained the round numeral *ten*, a proper granularity point allowing for imprecision/approximation. I take this additional "irrelevance" reason for not being more informative to be a separate implication numeral modifiers give rise to, which I will extensively deal with in chapter 7.

From the above we conclude that it is possible that once reading the numeral modifier in the target sentence, Westera & Brasoveanu's (2014) participants came up with additional reasons why the speaker used a numeral modifier and what that signals. Recall that a similar point was raised earlier when discussing McNabb & Penka's (2015) first experiment, which also involves a rich discourse context and a seemingly competent speaker. There too, we hypothesized that the observed results for the superlative test conditions could as well be because the speaker signals that she does not think it is relevant or does not want to share all the information/knowledge she has. I further noted that under such an assumption comparatives should perhaps behave in a similar way to superlatives. However, it was not possible to assess this assumption given McNabb & Penka's (2015) study, because they only tested superlative modifiers. Westera & Brasoveanu (2014) did include a comparative condition though, so let us evaluate their relevant results. Indeed, they do not find any difference between superlatives and comparatives in any region of the target sentence in the *what* condition in their first on-line experiment (no interaction effect, that is), nor in the region of the number (*ten*) or the first three spillover regions in the *how many* condition. Yet, Westera & Brasoveanu do find an interaction effect at the fourth region after the numeral (i.e., region *under* in (13)) in the *how many* condition. If this later effect is not to be connected with the processing of the modified numeral, one could infer based on the on-line

findings of the first experiment that, indeed, while incrementally reading the target sentences, participants were involved in reasonings about the speaker's utterance choices and signals other than that relating to speaker ignorance. That is, these reasonings were exhibited as slowdowns in regions following the numeral modifier in the *what* and *how many* QUD contexts, irrespectively of the numeral modifier type (note that the authors expected a difference between the two numeral modifiers in one of the two QUD conditions as far as speaker ignorance is concerned). If we are to maintain such a conclusion, we should perhaps acknowledge the habituation effect assumption Westera & Brasoveanu make to explain away the null findings of their second on-line experiment that included better-defined precise QUD conditions. Also, we shall assume that in the polar QUD condition no similar reasoning is triggered by the witness's use of a modified numeral, as this use is just made in the context of the extensive repetition of material from the relevant question without any communicative or informative purposes.

4.3 Wrap-up and further discussion

In the present chapter, I reviewed the existing experimental literature on the speaker ignorance inferences of numeral modifiers. We saw off-line studies involving a validity judgement task (Geurts & Nouwen, 2007; Geurts et al., 2010) that revealed a contrast between superlative and comparative modifiers to be attributed to the availability of ignorance inferences with the former modifiers and specifically to the semantic status of those inferences. As was pointed out, the relevant results can also be accommodated by a pragmatic account of speaker ignorance, as inference-validity judgment tasks test for any type of inference, be it semantic or pragmatic in nature. Furthermore, the use of a verification task argued to exclusively test for semantic inferences (Coppock & Brochhagen, 2013a) showed no difference between superlatives and comparatives, which did differ from the relevant control items, suggesting a non-semantic status of ignorance inferences, while being in line with a pragmatic one. Then a shift from methodologies with a binary (true/false) response paradigm to methodologies with a gradient response paradigm was reported. This shift was motivated by the assumption that the latter methodologies allow participants to distinguish among entailments, logical contradictions and pragmatic infelicities in their judgements, which has been experimentally demonstrated. Experimental investigations employing such methodologies aimed to explore the possibility that speaker ignorance inferences are pragmatic inferences. Participants now made their judgements on Likert scales of coherence (Cummins & Katsos, 2010; McNabb & Penka, 2015). The relevant experiments yielded results that suggest not only that ignorance effects are available with superlative as opposed to comparative modifiers, but also that they are pragmatic inferences, as the test conditions in question behave significantly different from the control condition of contradictions. However, the results in McNabb

& Penka (2015) can be interpreted as indicating that participants might have also been involved in other types of reasoning than just that deriving speaker ignorance.

Next, there were more direct manipulations testing for speaker ignorance (McNabb & Penka, 2015; Westera & Brasoveanu, 2014). McNabb & Penka manipulated the speaker's epistemic state in an off-line coherence judgement task (with a Likert scale) and Westera & Brasoveanu manipulated the QUD-dependence of speaker ignorance inferences in a self-paced reading task combined with a validity judgement task. Nevertheless, due to hidden confounding factors, as far as the former experiment is concerned we are only left with indirect evidence of speaker ignorance inferences and of their pragmatic status, while the latter experiment offers direct evidence of speaker ignorance inferences and their context-dependent status, resulting only from the validity judgement tasks; the on-line findings do not fully corroborate those findings. Last, Geurts et al.'s (2010) results of the on-line verification task provide no more than just indirect indications of speaker ignorance and its pragmatic status.

Having said that, all those studies make a significant contribution by revealing (mostly indirectly) that speaker ignorance inferences are available with superlative numeral modifiers and are not derived as entailments but rather as pragmatic inferences. The relevant findings, summarized above, disfavor semantic accounts like Geurts & Nouwen (2007) or Nouwen (2010), or Spector's (2015) grammatical account of speaker ignorance, and are compatible with a number of pragmatic (to a certain extent) proposals such as Büring (2008); Ciardelli et al. (2017); Cohen & Krifka (2014); Coppock & Brochhagen (2013b); Cummins & Katsos (2010); Kennedy (2015); Nouwen (2015); Schwarz (2016a); Spychalska (2015). It is further shown that comparative numeral modifiers can also trigger speaker ignorance as long as there is a precise question under discussion. This finding comes to terminate the longstanding assumption that only superlative modifiers—or class B modifiers in general, according to Nouwen (2010)—can give rise to speaker ignorance inferences.

Despite the substantial finding that speaker ignorance inferences of modified numerals are pragmatic in nature, we are still left with a considerable number of potential accounts, which each derives speaker ignorance via some pragmatic mechanism. What seems urgent now is to adjudicate between those remaining theoretical proposals. Those proposals comprise the neo-Gricean Quantity-based accounts of speaker ignorance, the speech-act accounts as well as the inquisitive semantics accounts (see relevant classification in chapter 2), with the latter two classes of accounts deriving a more obligatory type of speaker ignorance for superlative modifiers compared to the former. This obligatoriness is the result of combining a conventional basis for speaker ignorance with an appropriate pragmatic mechanism. Another open—related—issue concerns the direct investigation of the speaker ignorance inferences of numeral modifiers and of their time course in particular. As previously discussed, although Westera & Brasoveanu (2014) deliver important novel off-line data from their direct

investigation of speaker ignorance, which McNabb & Penka (2015) fail to do, their on-line findings are inconclusive. Note that contrary to the null reading-time results reported in Geurts et al. (2010), Westera & Brasoveanu do find significant effects in their self-paced reading experiment. However, they fail to replicate those effects in their second similar experiment. Moreover, there is a number of possible confounding factors that do not allow us to link the on-line findings of their first experiment to the off-line ones. Arguably, the most important of those is that the observed processing penalty could as well be due to readers' on-line reasoning about what the witness signals by using a modifier numeral, besides lack of knowledge, which is explicitly enforced by the judge's conclusion after the target sentence. Consequently, there is no solid evidence as to the real-time interpretation of superlative modifiers. Crucially, an experimental investigation of the real-time interpretation of superlative modifiers—which would specifically incorporate an appropriate and well-designed paradigm—can shed interesting light on the underlying processes responsible for the derivation of speaker ignorance inferences and thus ultimately help us differentiate among the various pragmatic proposals. In light of this urgent need, I set out to conduct an eye-tracking study aiming to bring to light the first direct evidence of the on-line accessing of speaker ignorance inferences. It was considered necessary to construct a design that primarily manipulates the speaker's epistemic state in such a way that the accounts differing in implicature obligatoriness will yield different predictions. This study consists of two eye-tracking reading experiments. Chapters 5 and 6 report on these experiments in great detail.

Furthermore, what I discussed above applies to comparative modifiers as well. That is, despite Westera & Brasoveanu's (2014) novel finding that comparative modifiers too can trigger speaker ignorance, due to the possible confounding factors in their first on-line experiment and the replication failure of the relevant findings, the following important issues remain unresolved: How are speaker ignorance effects of comparative modifiers derived and accessed in real time? Are the relevant processes different from those of superlative modifiers, and if so, in what way are they different? Resolving these issues will help us obtain a more concrete and complete idea of what an analysis should incorporate in order to account for the differences and similarities of superlative and comparative modifiers with respect to ignorance. The second eye-tracking reading experiment we carried out, presented in chapter 6, will also provide an answer to the aforementioned questions.

In the next two chapters (5 and 6), we will continue the study of speaker ignorance inferences, tackling experimentally all major issues left unsettled by the relevant empirical research up until now.

CHAPTER 5

The time course of speaker ignorance effects: Eye-tracking experiment 1

5.1 Introduction

In the previous chapter, we reviewed the experimental work that has recently developed and been conducted on numeral modifiers with the aim to investigate the availability of speaker ignorance effects as well as the extent to which those effects arise with various numeral modifiers. Employing methodologies that have extensively and effectively been used in other areas of the—also new and developing—field of experimental semantics/pragmatics, those studies further sought to answer the question whether ignorance inferences are semantic inferences or arise via pragmatic reasoning. Besides the variation of tasks, the studies also differed in whether they tested for ignorance directly or indirectly, with the majority of them using off-line techniques indirectly testing for ignorance inferences. Although, as we elaborately discussed, the various studies face certain issues, it is relatively clear by now that their findings indicate that the ignorance implication of superlative modifiers does not follow from their semantics, but it rather seems to have the status of a pragmatic inference. Thus, the experimental data appear to confirm the dominant theoretical view in the literature whereby speaker ignorance implications are generated from some pragmatic mechanism, which constitutes a substantial finding. However, the data in question do not really help us differentiate between the many available options in the pragmatic literature of numeral modifiers. The ultimate goal of the present chapter is to achieve exactly this.

More specifically, in the absence of direct investigations of ignorance ef-

fects yielding firm and conclusive results (see section 4.2.2), in this chapter, we directly probe ignorance effects and the time course thereof, measuring what happens in real time when interpreting a superlative modifier in a context where the information state of the speaker is manipulated (knowledgeable vs. ignorant). I do so by means of an eye-tracking reading experiment aiming to obtain insight into how and when ignorance effects arise and to examine to what extent such insights can inform the existing theoretical proposals as to the nature of the mechanism that is responsible for the effects in question.[1]

Section 5.2 explains why an eye-tracking study is appropriate for the purposes of our investigation. Section 5.3 presents the reading eye-tracking experiment we conducted, while section 5.4 discusses our findings and possible confounds thereof.

5.2 Why do reading eye-tracking?

As our aim is to investigate the incremental interpretation of superlative numeral modifiers in a context with an ignorant as opposed to a knowledgeable speaker and to obtain direct evidence of ignorance effects, the eye-tracking reading technique appeared to be the most suitable methodology. It is a sensitive method of detecting effects and processes that occur during reading, i.e., phenomena that hinder or even disrupt comprehension, from syntactic ambiguity to pragmatic violations. Specifically, it provides high temporal resolution, as it monitors eye-movements from millisecond to millisecond, as well as a variety of measures, thereby giving highly precise information about where readers look and for how long, and, by extension, about real-time sentence processing.

The underlying assumptions are that cognitive processes that are taxing and difficult for the processor or unfold over multiple steps interfere with comprehension, causing longer reading times, (higher probability of) re-reading of or regressing to previous words or parts of the sentence. Such processes triggered by fixating on a certain word might also extend to and affect the processing of subsequent fixated words, giving rise to the so-called spillover effects. Also, the timing and location of effects (relative to a critical word/region) can shed light on the nature of the possible challenging processes, although there is no strict linking between such characteristics and types of processes (Clifton, Staub & Rayner, 2007). It has been assumed that effects that occur when people first fixate on a certain region until they leave it can be associated with processes that occur in early stages of sentence processing, such as word recognition and structure building, while effects whereby people re-read or regress back to a certain region, effects that can also influence the reading of subsequent regions, are perhaps sensitive to higher level processes such as semantic integration and discourse processes (Boland, 2004). However, once again, those relations are not stringent or exhaustive and, thus, we should be careful with mapping early

[1] This experiment has also been reported in a paper that appeared in SALT 26 proceedings as Alexandropoulou, Dotlačil & Nouwen (2016).

and late effects (i.e, effects in so-called early and late measures, see section 5.3.6) onto early and late stages of cognitive processing, respectively, (see also Boston, Hale, Vasishth & Kliegl, 2011).

A major advantage of eye-tracking over other methodologies is that eye-tracking reading experiments allow for natural reading and for regressing to previous regions as well as re-reading certain regions, contrary, for instance, to self-paced reading experiments. In addition, eye-tracking testing allows participants to sit comfortably and read in a natural setting and environment, unlike EEG and fMRI experiments.

To end, the combination of the eye tracking reading methodology given its aforementioned advantages with the manipulation of the speaker's epistemic state in the context preceding the target sentence with the superlative numeral modifier will enable us to (i) look into the naturalness and compatibility of the use of the numeral modifier with the individual contexts, (ii) investigate the time course of accessing speaker ignorance inferences during the incremental interpretation of sentences with superlative modifiers, and (iii) reach a better understanding of the reasoning processes and mechanisms involved in the real-time interpretation of such occurrences of superlative modifiers.

5.3 Ignorance experiment 1

The present study was conducted in Dutch and consists of two pretests and the eye-tracking reading experiment.

5.3.1 Design

The numeral modifier we tested was the superlative *minstens*, which is the Dutch equivalent of *at least* (see Nouwen, 2010 re their parallel behavior). We looked at both unembedded and embedded occurrences of *minstens* in a context with a partially knowledgeable (ignorant) speaker as opposed to a fully knowledgeable (authoritative) speaker. So we had two manipulations:

(i) the speaker's epistemic state as set up by the context preceding the target sentence with *minstens* (Context manipulation): *speaker ignorance* vs. *speaker authority* context. This is similar to McNabb & Penka's (2015) context manipulation (in their second experiment, see section 4.2.2).

(ii) the main verb of the target sentence (Verb manipulation), as we included both *modal* (*moeten* 'must' / *willen* 'want to') and *non-modal* verbs (*hebben* 'to have' / *zijn* 'to be'), that is, creating embedded and unembedded occurrences of *minstens*, respectively.

Thus, the factors Context and Verb were manipulated in a 2×2 design. Below you see an example of a test item in all four resulting conditions. Texts consisted of an intro, followed by the context sentence with the Context manipulation,

followed in turn by the target sentence with the superlative modifier and the Verb manipulation; some texts would end with an outro (see details in section 5.3.5.2).

(1) **Example item**

Intro
Wesley heeft zijn eigen zaak waar hij met veel plezier tatoeages zet. Hij maakt er altijd echte kunstwerken van en het is er dan ook meestal erg druk. Hij wilde deze donderdag daarom hard doorwerken.
'Wesley runs his own tattoo parlor, which he enjoys a lot. He always makes real works of art and it is usually very busy. That's why he wanted to work hard all through Thursday.'

Context sentence

Speaker ignorance context
Ik weet niet precies hoe het met de drukte zat, maar ik heb wel een idee.
'I don't know exactly how busy it got, but I have got an impression.'

Speaker authority context
Ik weet precies hoe het met de drukte zat en daarom zal ik je erover vertellen.
'I know exactly how busy it got, and I'll tell you about it.'

Target sentence
Wesley $\left\{\begin{matrix}\text{moest}\\ \text{heeft}\end{matrix}\right\}$ *minstens zes mensen op donderdag* $\left\{\begin{matrix}\text{tatoeëeren}\\ \text{getatoeëerd}\end{matrix}\right\}$.
Wesley $\left\{\begin{matrix}\text{had to}\\ \text{has}\end{matrix}\right\}$ at least six people on thursday $\left\{\begin{matrix}\text{tattoo}\\ \text{tattooed}\end{matrix}\right\}$.
'Wesley had to tattoo/tattooed at least six people on Thursday.'

Outro
Hij heeft ook twee medewerkers in dienst die de kunst van het tatoeëeren van hem hebben geleerd.
'He also has two employees in his service that have learned the art of tattooing from him.'

In the following, I will elaborate on why we set up such a design and what this aimed to achieve.

Manipulation (i). Our primary aim concerns the first manipulation of our experiment (see (i) above). Our relevant context setup is similar in spirit to that used by Breheny et al. (2006), and Bergen & Grodner (2012) (see also Politzer-Ahles & Fiorentino, 2013, a.o.). To illustrate with Breheny et al.'s (2006) study, they aimed to directly measure what happens in real time when people interpret a scalar expression in two different types of contexts: Contexts

that bias comprehenders toward drawing the relevant scalar implicature versus contexts that do not have such a bias, but are rather merely compatible with the basic, lower-bound meaning of the scalar expression. They did so by means of a self-paced reading task and tested the Greek equivalents of the scalar expressions *or* and *some* (in different experiments). Below I exemplify with *some*, breaking down a translated example item into the context part with the context manipulation, the target sentence, and the continuation sentence.

(2) Example item (adapted from Breheny et al., 2006)

Context with bias: Mary asked John whether he intended to host all of his relatives in his tiny apartment. John replied that

Context without bias: Mary was surprised to see John cleaning his apartment and she asked the reason why. John told her that

Target sentence: he intended to host *some of* his relatives.
Continuation: The rest would stay in a nearby hotel.

As is evident, the **context with bias** provides a cue for drawing the *not all* scalar implicature, while no such cue is provided by the **context without bias**, hence fewer or no implicatures are expected to arise in the latter context condition compared to the former one. By analogy, in our experiment a speaker ignorance implicature is expected to arise in the target sentence in the ignorance contexts, which have an ignorance bias. On the other hand, no such implicature is expected in the authority contexts, because they do not have any bias and they are even contradictory with ignorance (while being compatible with the core meaning of the superlative modifier in the target sentence, i.e., $n \geq 6$).

Let me elaborate on our implementation of the above paradigm. First, the ignorant speaker in the ignorance context expresses that she does not have complete knowledge of some precise quantity (of people Wesley tattooed/had to tattoo in (1)), but she does have an impression, or opinion, or she can give an estimation. That is, our ignorant speaker appears to have partial knowledge. We did not want to have an ignorant speaker with complete ignorance in the relevant contexts, because it would be extremely odd and unnatural for such a speaker to utter any number in the following discourse, which could perhaps cause unwanted confounding delays during reading. Our ignorance contexts seem to entail speaker ignorance. Crucially though, the relevant implication is similar to the underspecified speaker ignorance interpretation (discussed in section 2.2.3), whereby the speaker lacks any beliefs about numbers in a certain range. That is, the speaker ignorance context sentence in (1) together with the basic meaning of the *at least*-modified numeral in the target sentence (semantically) expresses that the speaker does not know the exact number n in question, but knows that n is in the range $[6, ...)$, and this is all she knows; she lacks any beliefs about (the) values in that range.

As it will become useful and relevant for the following, it is essential to understand that the ignorance contexts yield a bias and they do not just instantiate what the ignorance implicature does via an inference. But why would the ignorance condition bias towards calculating an implicature if the context already contains the information conveyed by this implicature? The key is that the ignorance contexts do not contain the information that the specific ignorance implicature (also defined in section 2.2.3) expresses. Remember that specific speaker ignorance is the ignorance implicature that most of the pragmatic accounts derive (see section 2.2). According to such implicatures the speaker has (more) specific beliefs with respect to values in the range in question. Namely, in the example item in (1), on a specific ignorance reading, the speaker holds possible that $n = 6$ and she holds possible that $n > 6$. Importantly, there is experimental evidence providing insight into the type of implicature mechanism and of alternatives modified numerals are associated with, such that modified numerals like *at least n* involve an inference of a witness-possibility for the lowest number compatible with the modified numeral (Alexandropoulou et al., 2016; Mendia, 2016c, see also chapter 3). Specifically, Mendia (2016c), who studied experimentally speaker ignorance inferences, showed that in utterances like the target sentence in (1), $n = 6$ necessarily constitutes an epistemic possibility for the ignorant speaker, going against Coppock & Brochhagen's (2013b) and Nouwen's (2015) anti-specificity-based-only relevant accounts. As is obvious, such an inference is what primarily defines what I have called specific speaker ignorance implicature.

Therefore, given this evidence, the speaker ignorance contexts (providing ignorance cues) in the present experiment will favor the calculation of a specific ignorance implicature in the target sentence, irrespective of the main verb used in the sentence (see further below for ignorance interpretations of embedded *at least*). Remarkably, besides Breheny et al. (2006), there is a big number of studies on scalar expressions showing that people draw the relevant scalar implicature robustly when this is contextually justified (Bergen & Grodner, 2012; Degen & Tanenhaus, 2014; Hartshorne, Liem Azar, Snedeker & Kim, 2014; Lewis, 2013; Politzer-Ahles & Fiorentino, 2013; Politzer-Ahles & Gwilliams, 2015).

More importantly, such implicatures are not expected to arise in the authority context condition. Given our setup, the authoritative speaker always conveys that she has complete knowledge of some precise quantity, signaling it is relevant in the context, to which the aforementioned ignorance implicatures are diametrically opposite. This explains why we expect authority contexts to be just compatible with the basic meaning of *at least* in the target sentence that follows, again, in either verb condition. Also, this is in line with recent experimental findings by Cremers et al. (p.c.) revealing that *at least* can (highly) appropriately be used to refer to a certain quantity in scenarios where the speaker is without doubt fully knowledgeable of the precise quantity.

Manipulation (ii). Speaker ignorance effects can arise not only when *at least* occurs unembedded, but also when it is embedded in the scope of a universal modal, as McNabb & Penka (2015) showed experimentally (see section 4.2.2). In the target sentence of the example item above, when *minstens* ('at least') is embedded under the universal modal *moest* ('had to'), the relevant speaker ignorance reading is that the speaker does not know the exact (personal) requirement/desire of Wesley's, that is, exactly how many people Wesley had to tattoo on Thursday. The dominant take in the literature is that speaker ignorance inferences in such environments are triggered as the result of the modified numeral taking scope over the modal operator (see Blok, 2018 for a different account). By including both embedded and unembedded occurrences of the superlative modifier in the present experiment, we aim to test whether the time course of accessing speaker ignorance inferences during the incremental interpretation of the respective sentences would be different in any respect; for instance, in case speaker ignorance effects are more robust in the non-embedding as opposed to the embedding environments, but, more importantly, whether traces of wide scope movement of *at least* with respect to the relevant modal operator could be detected in the latter environments by the eye-tracking means.

What the various theoretical accounts have in common is that they posit and require a wide-scope LF in order for ignorance to arise in embedding environments. Specifically, an ignorance interpretation comes about due to the following configuration for each account (I am including those accounts that explicitly discuss such embedding environments):

(i) when *at least n* outscopes the present universal modal operator $\Box$ according to the neo-Gricean accounts (Büring, 2008; Kennedy, 2015; Nouwen, 2015). Then the standard Quantity-based reasoning (see section 2.2.2 and also 2.3.1.2) applies to the alternatives (introduced now by $\Box$, schematically: $\Box[6]$ and $\Box[7, ...)$ for the target sentence in (1)). This gives rise to the primary implicatures: *the speaker does not believe that Wesley had to tattoo exactly 6 people* (i.e., $\neg\Box_{Bel}\Box[6]$) and *the speaker does not believe that Wesley had to tattoo more than 6 people* (i.e., $\neg\Box_{Bel}\Box[7, ...)$).

(ii) when *at least*, which quantifies over alternative propositions/possibilities in Coppock & Brochhagen's (2013b) inquisitive semantics framework, is interpreted above the universal modal. Given that and the Interactive Sincerity maxim, this results in a speaker's epistemic state that consists of more than one possibility as far as the lower bound of the relevant range is concerned. For the target sentence in (1), the (ignorant) speaker entertains the following epistemic possibilities: *Wesley had to tattoo 6 people on Thursday*, *Wesley had to tattoo 7 people on Thursday*, *Wesley had to tattoo 8 people on Thursday*, etc.

(iii) when the epistemic modal quantifier introduced by the semantics of *at least* takes scope over a universal deontic modal, as in the semantic ap-

proach by Geurts & Nouwen (2007). Given the modal meaning they assign to *at least*, the resulting compositional semantics gives the following for the target sentence in (1): *The speaker believes that Wesley had to tattoo six people on Thursday and she considers it possible that Wesley had to tattoo more than six people on Thursday.*

(iv) when the GRANT illocutionary operator *at least* expresses under Cohen & Krifka's (2014) account takes wide scope with respect to the universal modal. I.e., for the target sentence in (1) the basic meaning is that the speaker denies that Wesley had to tattoo fewer than six people, while the implicated speaker ignorance about the exact number of people amounts to: for any $m \geq 6$ the speaker considers it possible that Wesley had to tattoo exactly m people.

As speaker ignorance contexts entail ignorance and bias toward a speaker ignorance inference, a wide-scope configuration for *at least* relative to the present modal will be facilitated in such contexts in the modal Verb condition.

The aforementioned accounts further derive a narrow-scope reading for *at least*, which is an authoritative/speaker knowledgeability reading, whereby *at least n* is interpreted in the scope of a present universal modal. This is actually the most prominent interpretation of the interaction of *at least* with a universal modal, when a relevant sentence is uttered out of the blue. To illustrate for the modal version of the target sentence in (1), the narrow-scope reading in question is: the speaker knows Wesley's precise desire/personal requirement, i.e., that his personal requirement would be fulfilled as long as he tattooed m-many people on Thursday, where $m \geq 6$. This scope configuration will be favored when the modal target sentence follows a speaker authority context sentence, which is a knowledgeability context, incompatible with speaker ignorance, providing no specific cue whatsoever.

Having explained what our manipulations aim to test, we are ready to go through the predictions that arise from the existing theoretical accounts when taking into account the context setup of our experiment.

5.3.2 Predictions of theoretical accounts

In this section, I will extract predictions from various theoretical accounts in the light of the context setup we employ in the present experiment. Although the semantic accounts of speaker ignorance have already been disqualified by the available experimental findings, as demonstrated in chapter 4, I will still consider in the following the predictions arising from those accounts. It is important to evaluate these predictions against the present paradigm and see whether the findings of the current study are in line with the existing ones with respect to the accounts under consideration.

Below, I will start with the predictions concerning the manipulation of the speaker's epistemic state in the context in interaction with the unembedded

uses of *at least* and continue with those concerning that manipulation in combination with the embedded uses of *at least*, when in the scope of a universal modal operator. In the end, based on the discussed predictions of the different theoretical accounts, I will formulate broader testable hypotheses and the corresponding predictions given the design of experiment 1.

5.3.2.1 Predictions re unembedded uses of *at least*

What will be crucial for the relevant predictions is that some accounts derive obligatory, context-independent speaker ignorance inferences, by virtue of the truth-conditional meaning of a sentence with *at least* or by virtue of a certain inescapable (pragmatic) requirement associated with such an utterance, while others posit no restriction on *at least* that would impose an obligatory, context-independent ignorance inference in non-embedding environments (cf. relevant remarks on ignorance obligatoriness throughout section 2.2 in relation to the individual theoretical accounts discussed).

Let's start with the former type of accounts. Geurts & Nouwen (2007) assume a(n epistemic) modal component in the lexical meaning of *at least*, so any (unembedded) use of *at least* is expected to have a(n epistemic) modal implication. Hence, the following are predicted: (i) compatibility of the target sentence with the preceding context when it biases towards speaker ignorance, and (ii) incompatibility of the target sentence with the preceding context when it sets up a knowledgeable speaker, due to the contradiction of the speaker's epistemic state as set up by the context and that signaled by the semantics of *at least*. The exact same predictions hold for Nouwen's (2010) account, which assumes that unembedded occurrences of *at least* are licensed by a silent existential modal operator, which is responsible for the ignorance implication. Thus, although such occurrences of *at least* will be compatible with a speaker ignorance context, we expect a clash between the target sentence and a context sentence with an authoritative/knowledgeable speaker: The logical form that corresponds to the target sentence has a(n obligatory) speaker ignorance implication contradicting with the relevant context. In a similar way, Spector (2015), who adopts a grammatical approach to implicatures, derives obligatory Quantity-based speaker ignorance implicatures with unembedded uses of *at least* stemming from the obligatory application of an LF exhaustification operator. Given that, a contradiction is expected between the compositionally derived obligatory speaker ignorance implication of the target sentence and speaker authority contexts, whereas semantic compatibility and coherence obtain with combinations with a speaker ignorance context.

We shall now move to more hybrid accounts of speaker ignorance, which derive speaker ignorance as an implicature, but crucially establish both a conventional and a conversational (pragmatic) basis for it. Ciardelli et al.'s (2017) and Coppock & Brochhagen's (2013b) accounts were characterized as such in section 2.2.3. These accounts propose an alternative-generating (or else possibility-introducing) semantics for *at least* and derive an obligatory speaker ignorance

inference via the independently motivated conversational maxim of Inquisitive Sincerity or of Interactive Sincerity, respectively. Those Quality maxims require that a speaker should not use *at least* if she already knows how to resolve the issue an utterance with *at least* expresses. Given that Quality implicatures are robust and inescapable implications (see relevant discussion in section 2.2.3) and that they are context-independent according to the accounts under consideration, the following is predicted: In the authority Context condition the target sentence with *at least* will be incompatible with the preceding context that reveals that the speaker knows which of the possibilities/alternative propositions holds, namely, the precise number in question (e.g., exact number of people Wesley tattooed/had to tattoo in (1)). This incompatibility arises due to the violation of the relevant sincerity maxims, which are, in contrast, obeyed by the ignorant speaker in the ignorance context condition.

Spychalska's (2015) account, too, can be considered a hybrid account, which derives a kind of obligatory speaker ignorance as well. Spychalska (2015) makes a distinction between truth- and assertibility conditions and actually takes speaker ignorance to be part of the assertibility conditions of *at least*. While, roughly speaking, *at least* n is false if the quantity under discussion is smaller than n and true if it is exactly n or greater, the assertibility conditions for *at least* n require that the speaker considers it possible that the quantity in question is n and considers it possible that it is greater than n. Both the truth- and the assertibility conditions are met by the ignorant speaker in the ignorance contexts, whereas this is not the case in the authority contexts for the assertibility conditions. The speaker's epistemic state as defined by these specific contexts contradicts the speaker's beliefs as prescribed by the assertibility conditions of *at least* used in the target sentence that always follows. Given the robustness of assertibility conditions (see discussion in section 2.2.4 of chapter 2), this contradiction is expected to be expressed as an infelicity of the target sentence given the preceding context sentence in the authority Context condition.

The last account to be considered as hybrid is that by Cohen & Krifka (2014), who seem to derive an obligatory speaker ignorance implicature with unembedded occurrences of *at least*, as the derivation of the implicature is necessary for accounting for the truth of the *at least*-sentence. To remind you, on this account a speaker that produces an utterance with *at least* n signals that for any $m \geq n$ she is not committed to the assertion $\neg[m]$, that is, the speaker leaves it open whether $[n]$ is false, $[n+1]$ is false, $[n+2]$ is false, etc. As a result, speaker (total) ignorance is drawn as a possible reason why the speaker did not make a stronger utterance/commitment that excludes assertions of the propositions that m-many individuals satisfy a certain property, for each $m \geq n$. This speaker ignorance implicature is necessarily triggered in the ignorance Context condition and it is also necessarily drawn in the speaker authority Context condition. The latter results in a clash, because the knowledgeable speaker is committed to the assertion that *exactly* n individuals satisfy the property in question.

In the foregoing we saw accounts that derive a mandatory type of speaker ignorance inferences for unembedded uses of *at least*, which, while being in accordance with speaker ignorance contexts, come into conflict with the speaker authority contexts. Next, we will see accounts that do not assume any strict restriction on *at least* and its use that would require obligatory, context-independent speaker ignorance. Hence, speaker ignorance is expected to arise when it is supported by the preceding context, e.g., by our speaker ignorance context in (1). This means that no ignorance implication is predicted to arise in the contexts that do not provide any specific cue for a certain inference and that reveal a knowledgeable speaker, cf. the speaker authority context in (1). But let us see why.

The accounts under consideration are the neo-Gricean accounts by Büring (2008), Cummins & Katsos (2010), Kennedy (2015), Nouwen (2015), and Schwarz (2016a), which derive ignorance effects with *at least n* as (primary) Quantity implicatures (*the speaker does not believe* $[n]$ *to be true and does not believe* $[n+1, ...)$ *to be true*). Importantly, these accounts do not make a(ny) specification that would necessitate the derivation of an ignorance inference with every unembedded occurrence of *at least* and also regardless of context. Furthermore, note that there is a great deal of studies corroborating the optional and context-dependent character of Quantity implicatures (see, e.g., Bergen & Grodner, 2012; Breheny et al., 2006; Degen & Tanenhaus, 2014, a.o.). Consequently, we expect the following given our context setup: A (specific) speaker ignorance inference will become available in the ignorance Context condition, which provides a strong cue to such an implicature, while the opposite is expected in the authority Context condition; that is, a speaker ignorance implicature is less likely to arise in a context that has no specific bias and the relevant speaker reveals that she is fully knowledgeable. This means that the target sentence with *at least* will be equally felicitous when following a speaker ignorance or a speaker authority context sentence.

I should note here that Nouwen's (2015) account is not specific as to when or how robustly the pragmatic reasoning about the alternatives having to do with the domain anti-specificity presupposition applies. This has already been remarked in section 2.2.2. Neither Alonso-Ovalle & Menéndez-Benito (2010), which Nouwen's (2015) proposal is based on, makes such a specification. How should we think of Nouwen's analysis then? From now on I will assume that the relevant pragmatic reasoning is a Manner-like reasoning, whereby the listener wonders why the speaker used an expression that presupposes a bigger domain rather than an expression with a singleton domain. It has been shown experimentally that Manner implicatures are optional, and world-knowledge- and context-dependent (Bott, Frisson & Murphy, 2009). Taking this into account, we take the reasoning in question in Nouwen's (2015) account to not be obligatorily initiated in the authority or the ignorance Context condition of our experiment. That is, in the latter condition the context sentence will prompt the reader to draw the Manner-based, as well as the additional Quantity-based, reasoning, and, thus, the relevant specific ignorance implicature, while in the

former condition no implicature will be derived and the target sentence will be consistent with the preceding context sentence.

As the crucial aspect of the accounts discussed above is the obligatoriness of speaker ignorance implications of *at least*, below I group the relevant predictions of the individual accounts based on this aspect. Specifically, Table 5.1 subsumes the theoretical accounts under two broader hypotheses in relation to the obligatoriness of the ignorance implication they derive and includes the ensuing predictions as regards the unembedded uses of *at least* when considering the context setup of experiment 1.

Hypotheses	**Predictions re unembedded *at least***
TH1 Obligatory ignorance	**Target** compatible with **ignorance context** **Target** incompatible with **authority context** (CCR, CKR, CB, GN, N10, SPEC, SP)
TH2 Non-obligatory ignorance	**Target** compatible with **ignorance context** **Target** compatible with **authority context** (B, CK, K, N15, S)

Table 5.1: Hypotheses re obligatoriness of ignorance implication and predictions given the design of eye-tracking experiment 1. B stands for Büring (2008), CCR for Ciardelli et al. (2017), CKR for Cohen & Krifka (2014), CB for Coppock & Brochhagen (2013b), CK for Cummins & Katsos (2010), GN for Geurts & Nouwen (2007), K for Kennedy (2015), N10 for Nouwen (2010), N15 for Nouwen (2015), S for Schwarz (2016a), SPEC for Spector (2015), and SP for Spychalska (2015).

5.3.2.2 Predictions re embedded *at least*

As already explained, our context setup determines which scope configuration will be favored in the modal Verb condition of the target sentence (in each Context condition), although a narrow-scope interpretation of the superlative modifier might be the most salient reading of the target sentence. So given that the wide-scope ignorance interpretation is favored by the ignorance context, while the narrow-scope authoritative reading is favored by the authority context, the individual theoretical accounts derive these readings by their own means (see section 5.3.1), resulting in an interpretation of the target sentence that is compatible with the information conveyed by the preceding context sentence in both Context conditions.

Before moving to the processing predictions, I would like to note that the narrow-scope configuration of *at least* relative to the present modal operator has been argued to further trigger the so-called *modal variation* implication (recall from section 2.3 in chapter 2, and also chapter 3), such that according to the

speaker in (1), repeated below as (3), Wesley would be satisfied with tattooing six people on Thursday and he would be satisfied with tattooing seven people on Thursday, and ... with tattooing eight people on Thursday, etc. (Buccola & Haida, 2017; Büring, 2008; Coppock & Brochhagen, 2013b; Nouwen, 2015, among others).

(3) **Target sentence of example item**

Wesley moest minstens zes mensen op donderdag tatoeëeren.
Wesley had to at least six people on Thursday tattoo
'Wesley had to tattoo at least six people on Thursday.'

Most accounts derive this implication via pragmatic means. Importantly and as already highlighted, the speaker authority contexts of the present experiment do not provide any cue for a particular inference and certainly no cue for a variation of possibilities or of possible satisfactory scenarios according to the main character of the relevant texts. The only cue these contexts provide is that the speaker is knowledgeable of the minimum requirement/desire in question. Therefore, it is predicted that, if at all, due to lack of context support the modal variation implication will (at the very best) arise in the condition under consideration less robustly, and surely way less robustly compared to the ignorance inference prompted by the corresponding ignorance contexts.

5.3.3 Processing predictions

There are no explicit predictions as far as the processing of numeral modifiers and their incremental interpretation are concerned. The vast majority of the existing accounts consist in theoretical proposals with no direct processing implications. Nonetheless, there exist a few proposals incorporating a part that relates to processing. Cummins & Katsos (2010), for example, claim that superlative modifiers are *more difficult to process at a psychological level* (p. 279) compared to comparative modifiers, because the former have a disjunctive meaning (e.g., $= n$ *or* $> n$ for *at least* n), which is psychologically more complex than either of the disjuncts. Geurts et al. (2010) too, who embrace Geurts & Nouwen's (2007) modal account, maintain that because of their modal semantics superlative modifiers are harder to process compared to comparatives, which do not have such a semantics.

Geurts et al. (2010) justify their processing statement by means of the findings of their on-line verification task (discussed in section 4.2.1). The task showed that superlative modifiers as opposed to comparative modifiers (or the baseline *exactly* modifier condition) delay verification judgements of the sentences containing them (against a situation), though no difference was found in the (total) reading times of those sentences. As I also remarked in section 4.2.1, although the authors conclude that their findings indicate that superlative modifiers are (semantically) more complex, we can draw no firm conclusion with respect to their on-line interpretation. That is, one cannot be sure what

this delay in decision times should be associated with, especially given that this does not manifest itself in the reported reading times of the relevant sentences. Thus, it is very unclear why the modal interpretation of superlative modifiers would only occur in and affect the decision times.

Cummins & Katsos (2010) make a similar point and in order to justify their own processing assumption they conduct a similar experiment to that by Geurts et al.'s (2010) described above making the following modifications: They replace the numeral modifiers (*exactly, more than, less than, at least, at most*) with comparison operators (=, >, <, ≥, ≤), where the latter two operators are disjunctive consisting of the simpler operators (=, > and =, <, respectively). They find the same results as Geurts et al. (2010): The disjunctive operators delay the decision procedure compared to the respective simple operators, similarly to superlative modifiers in Geurts et al. as compared to the comparative and *exactly* conditions. Cummins & Katsos conclude that the observed effect has nothing to do with the supposed lexical modality of superlative modifiers, but it should rather be due to their disjunctive interpretation, which they do share with the operators ≥ and ≤.

However, we should be cautious with such a conclusion. The mean decision times for the disjunctive conditions in Cummins & Katsos (2010) (≥: 1110, ≤: 1131) are much shorter than the corresponding superlative conditions in Geurts et al. (2010) (*at least*: 1559, *at most*: 1982), while the difference between the = and *exactly* baseline conditions is not that big (982 and 1114, respectively). Even if we assume that (part of) the difference is due to Geurts et al.'s (2010) participants converting the linguistic expressions to the mathematical expressions or just due to processing spillover because of the length difference, still, it seems that a bigger complexity is possibly associated with superlative modifiers compared to the mathematical operators. All in all, we infer that although the two accounts in question include a processing profile on top of their theoretical one, the (real-time) processing evidence reported lends them very little support.

Neither can Westera & Brasoveanu's (2014) on-line findings be used in favor of the processing proposal of the accounts in question, which posit that superlative modifiers are psychologically more complex due to their modal (Geurts et al., 2010) or disjunctive (Cummins & Katsos, 2010) meaning. In their first self-paced reading task, Westera & Brasoveanu (2014) find that participants slow down when reading the region of the numeral that is being modified and at subsequent regions, but they do not find a difference between superlative and comparative modifiers, as Cummins & Katsos (2010) and Geurts et al. (2010) would predict. They only find a slowdown in the superlative condition five regions after the numeral modifier. Westera & Brasoveanu (2014) take the relevant effects to be due to possible silent intonational effects during reading or due to the costly on-line calculation of speaker ignorance inferences via pragmatic reasoning. Regarding the latter explanation, they back it up by making a link between the attested on-line effects and the ignorance interpretation rates they obtain in the off-line judgement task. However, as I have already discussed

(see section 4.2.2), such a link is not warranted, as we cannot be sure whether the participants generated ignorance inferences on-line in the target sentence or whether they only did so when explicitly prompted (by the judge's conclusion), after reading the target sentence, to adopt an ignorance interpretation of the target sentence. In other words, it is not necessary to associate the ignorance interpretation judgements participants made *after* reading the target sentence with how they processed the target sentence during self-paced reading. So this questions the validity of an on-line-implicature-calculation explanation of the attested processing effects. In any case, in a second similar on-line experiment Westera & Brasoveanu failed to replicate what they found about the incremental interpretation of sentences with numeral modifiers, while this time the experiment they conducted was better controlled and with an improved design. To conclude, the little processing predictions available in the literature of numeral modifiers seem to be disfavored by any existing conclusive experimental evidence.

Since there is no concrete processing profile for superlative modifiers in the literature, given the limited availability of (successful) predictions and definitive empirical evidence, we set out to investigate from scratch the time course (i) of interpreting the superlative modifier *at least* in a manipulated context and (ii) of accessing speaker ignorance inferences. Crucially, we will maintain a neutral approach to the processing of the superlative numeral modifier. That is to say, we will consider all possible processing implications that the predictions of the theoretical accounts from the previous section (section 5.3.2) could have. We will consider both the possibility that the on-line interpretation (of a felicitous occurrence) of the modified numeral may be associated with a processing difficulty and the possibility that it may be rapidly generated. We will furthermore consider the possibility that incompatibility of a speaker ignorance interpretation with the context, as predicted by a number of accounts for the use of unembedded *at least* in authority contexts, may (or may not) disrupt reading. All such possibilities will be put on the table on reasonable grounds and without taking sides, but rather letting the data speak for themselves. Every hypothesis to be discussed will be given a label for ease of reference. In what follows, we will proceed in a way parallel to the previous section.

5.3.3.1 Processing predictions re unembedded *at least*

Let us start with the predictions for the speaker ignorance Context condition, which is a point where all theoretical accounts seem to converge. It has been argued that a speaker ignorance inference will be triggered once people read the unembedded modified numeral phrase. How will the access to the ignorance inference be manifested on-line?

If speaker ignorance is derived as a semantic entailment (Geurts & Nouwen, 2007; Nouwen, 2010; Spector, 2015), in general we would expect no processing cost to be induced (hypothesis PH1) unless we assume that this inference is the result of a special process, which is different from the regular composi-

tional processes (PH2). That is, something should have been exceptional and special about the semantic modal component of *at least* in Geurts & Nouwen's (2007) account, the grammatical operation (i.e., exhaustification operation) employed by Spector (2015) to derive speaker ignorance, or the combinatorics of the silent existential modal operator with unembedded occurrences of *at least* under Nouwen's (2010) account. In fact, in the latter case, the insertion of the existential modal that derives the most common implication of superlative modifiers is motivated by Nouwen as a last resort strategy (see section 2.2.1). Such a strategy then can perhaps be assumed to incur some extra processing cost.

What should we consider for the pragmatic accounts of speaker ignorance, such as Büring (2008); Ciardelli et al. (2017); Cohen & Krifka (2014); Coppock & Brochhagen (2013b); Cummins & Katsos (2010); Kennedy (2015); Nouwen (2015); Schwarz (2016a), and Spychalska (2015), according to which a speaker ignorance implicature is predicted to arise in the relevant contexts? What are the processing predictions in that case? Given our agnostic approach, we shall assume the following: (i) If the pragmatic process involved is an effortful process (hypothesis PH4), as a large number of studies has shown for the process involved in the computation of scalar implicatures (Bott et al., 2012; Bott & Noveck, 2004; Breheny et al., 2006; Chevalier, Noveck, Nazir, Bott, Lanzetti & Sperber, 2008; Huang & Snedeker, 2009; Panizza, Chierchia & Clifton, 2009), we expect to find a processing penalty at the region of the modified numeral (or later) in the ignorance Context condition. Cf., for instance, a similar effect at the region of the scalar expression in the reading studies by Breheny et al. (2006) and Panizza et al. (2009) attested in the contexts facilitating the relevant scalar implicature. (ii) If the pragmatic process responsible for the derivation of speaker ignorance does not come with an extra cost but occurs rapidly (hypothesis PH3), as other similar studies have demonstrated for scalar implicatures (Breheny et al., 2013; Foppolo & Marelli, 2017; Grodner et al., 2010, a.o.), no processing penalty is expected in the ignorance Context condition.

Things get more interesting and divergent once we turn to the predictions concerning the speaker authority Context condition. This is the condition where most of the theoretical accounts vary, as also illustrated in section 5.3.2.1. We shall first consider those accounts that derive obligatory speaker ignorance inferences with unembedded occurrences of *at least* that predict some incompatibility of the target sentence with the preceding authority context. This incompatibility might arise because of the semantic inconsistency or contradiction of the speaker's assertion in the target sentence with what she states in the preceding context (Cohen & Krifka, 2014; Geurts & Nouwen, 2007; Nouwen, 2010; Spector, 2015), or because of the violation of some fundamental pragmatic maxim (Ciardelli et al., 2017; Coppock & Brochhagen, 2013b) or condition on the use of *at least* (assertibility condition in Spychalska, 2015) and the obligatorily derived ignorance implicature associated with the superlative modifier in the target sentence. I will refer to the former case as *semantic violation/contradiction* and to the latter as *inconsistency due to obligatory or*

hard-to-cancel implicature.

If a contradiction incurs a processing cost (hypothesis PH6), as it has been found, for instance, for the use of a word that is semantically anomalous given the wider discourse (van Berkum, Hagoort & Brown, 1999; van Berkum, Zwitserlood, Hagoort & Brown, 2003), we similarly expect a processing penalty and disruption of reading at the region of the modified numeral (or the spillover regions), whose integration to the representation should result in (self-)contradiction according to Cohen & Krifka (2014); Geurts & Nouwen (2007); Nouwen (2010); Spector (2015). If, on the other hand, such contradiction cases are not effortful (hypothesis PH7), which would be particularly unexpected and counterintuitive, we shall find no relevant effect in the target sentence.

Furthermore, if we assume that the resulting incompatibility between target and context sentences due to hard-to-cancel implicature (cf. inquisitive sincerity-based ignorance implicature in Ciardelli et al., 2017, interactive sincerity-based implicature in Coppock & Brochhagen, 2013b, assertibility-based implicature in Spychalska, 2015) is costly and disrupts comprehension (hypothesis PH8), an effect should occur at the modified numeral phrase in the target sentence. Whether or not the derivation of ignorance inferences is a costly process, this effect should arise as an interaction effect between Context and Verb factors, (i.e., non-modal Verb+authority Context).

If the arising inconsistency due to hard-to-cancel implicature (again as in Ciardelli et al., 2017; Coppock & Brochhagen, 2013b; Spychalska, 2015) is not a taxing process (PH9) though, no effect should be found in the target sentence in the authority Context condition. Hence, in that case we expect only the effect for the speaker ignorance contexts, if the on-line derivation of ignorance effects is costly, or no effect whatsoever, if it is not.

Next, given that no incompatibility is predicted in the authority Context condition by the neo-Gricean accounts, which assume no restriction on *at least* that would impose obligatory, context-independent ignorance inferences (Büring, 2008; Kennedy, 2015; Nouwen, 2015; Schwarz, 2016a, see also relevant prediction given TH2 on Table 5.9), no effect on eye-tracking measures is expected in the target sentence with unembedded *at least*.

Lastly, Cummins & Katsos (2010), who explicitly take superlative modifiers to be psychologically more complex because of their disjunctive status (hypothesis PH5), predict a processing penalty at the superlative modified numeral phrase across Context conditions.

Table 5.2 summarizes the processing hypotheses we considered above and the corresponding predictions given each account for ease of reference in the following. Once again, I would like to stress that these are tentative hypotheses about the real-time processing of the numeral modifier *at least* put forth in the absence of a processing theory and given the abundance of theoretical accounts. I do not intend to imply in any way, for instance, that accounts that include a standard Gricean machinery yield implications about the cognitive processes involved in the observance or violation of some Gricean maxim. That is to say,

we should definitely escape such an unfortunate misconception about Gricean theory (see more on this in Geurts & Rubio-Fernández, 2015) or any other type of theory, which happens quite often.

Hypotheses	**Predictions** re unembedded *at least*
PH1 Rapid derivation of	No precessing penalty
PH2 Costly derivation of *semantic* implications: • due to lexical semantics • due to last-resort insertion of covert existential modal • due to obligatory LF exhaustification operation	Processing penalty at *at least* phrase in **ignorance/authority contexts** (GN) (N10) (Spec)
PH3 Rapid derivation of	No precessing penalty
PH4 Costly derivation of *pragmatic* implications: • via Quality maxim • via assertibility conditions • via Quantity maxim	Processing penalty at *at least* phrase in **ignorance contexts** (CB, CCR) (Sp) (B, CKr, K, N15, S)
PH5 *At least* is psychologically complex due to disjunctive status	Processing penalty at *at least* phrase in **ignorance/authority contexts** (CK)
PH6 Semantic violation/contradiction is taxing **PH7** Semantic violation/contradiction is not taxing	Processing penalty at *at least* phrase in **authority contexts** No processing penalty (GN, N10, Spec, CKr)
PH8 Inconsistency due to hard-to-cancel implicature is taxing **PH9** Inconsistency due to hard-to-cancel implicature is not taxing	Processing penalty at *at least* phrase in **authority contexts** No processing penalty (CB, CCR, Sp)

Table 5.2: Processing hypotheses and predictions re unembedded occurrences of *at least*. B stands for Büring (2008), CCR for Ciardelli et al. (2017), CKr for Cohen & Krifka (2014), CB for Coppock & Brochhagen (2013b), CK for Cummins & Katsos (2010), GN for Geurts & Nouwen (2007), K for Kennedy (2015), N10 for Nouwen (2010), N15 for Nouwen (2015), S for Schwarz (2016a), Spec for Spector (2015), and Sp for Spychalska (2015).

5.3.3.2 Processing predictions re embedded *at least*

For the embedded occurrences of *at least* in a speaker ignorance context the predictions for each account are the same as for the unembedded occurrences

of *at least* discussed in the previous section. For the speaker ignorance to arise though yet another process is required to take place first, and that is the covert wide-scope movement of *at least* relative to the present modal operator. This is the scope configuration that triggers speaker ignorance.

Covert quantifier movement has been attested on-line and found to incur a processing cost during on-line comprehension (Hackl, Koster-Hale & Varvoutis, 2012, but see Jacobson & Gibson, 2014; Gibson, Piantadosi & Levy, 2017; Szabolcsi, 2014 for discussion). If this is so (hypothesis PH10), it is predicted that a greater processing penalty will be attested at the region of the modified numeral or later in the modal version as opposed to the non-modal version of the target sentence in the ignorance Context condition. Hence, an interaction of the Context and Verb factors is to be expected according to the accounts that derive wide-scope ignorance inferences for embedded occurrences of *at least* (Büring, 2008; Cohen & Krifka, 2014; Coppock & Brochhagen, 2013b; Geurts & Nouwen, 2007; Kennedy, 2015; Nouwen, 2015). If this is not so (hypothesis PH11), we only expect the effect due to the on-line derivation of speaker ignorance, assuming that this is a costly operation, while no effect at all is expected assuming otherwise.

Note that nothing special is expected to happen in the target sentence when it has a modal main verb and is preceded by a speaker authority context. Such a context is incompatible with speaker ignorance and gives no cue whatsoever besides signaling that the speaker is knowledgeable, thereby favoring a narrow scope reading of *at least* such that the speaker is fully informed about the precise requirement/desire expressed by the target sentence. Hence, no ignorance or other inferences are biased or predicted to arise in this condition. As commented in section 5.3.2.2, if modal variation inferences arise in the target sentence, although not explicitly prompted by the context, this would be expected to happen to a smaller degree compared to the speaker ignorance inferences that are being clearly encouraged by the relevant context in the corresponding ignorance condition. Table 5.3 presents the basic hypotheses to be tested with respect to the processing of the embedded occurrences of *at least*.

Hypotheses		**Predictions re embedded *at least***
PH10	Wide-scope quantifier movement is costly	Processing penalty at *at least* phrase in modal condition of **ignorance contexts**
PH11	Wide-scope quantifier movement is not a costly process	No interaction effect in **target sentence** due to movement

Table 5.3: Processing hypotheses and predictions re embedded occurrences of *at least*.

As you notice, in the predictions put forth throughout the current section, there is no further specification with respect to the various eye movement measures. One could claim that since all our hypotheses concern high-level processes, i.e., semantic and pragmatic processes, the predicted effects will appear later in the eye movement record and possibly in post-critical regions (see, e.g., Fukumura & van Gompel's (2017) findings on processing of pragmatic maxim violation). However, there is also a considerable number of studies specifically showing that semantic and pragmatic anomaly affect the eye movement record early on once the critical word is read (e.g., Murray & Rowan, 1998; see also Clifton et al., 2007 and references therein), but, as Clifton et al. (2007) also point out, this could *reflect the magnitude of the processing disruption* (p. 365), which possibly has to do with the nature of the semantic/pragmatic process that takes place. Perhaps a pragmatic anomaly (e.g., pragmatic implausibility of a word) is more catastrophic than a pragmatic maxim violation. Since the findings in the eye-tracking literature are split as far as high-level factors are concerned, we should rather avoid making more specific predictions.

5.3.4 Pretests

Before running the eye-tracking experiment we conducted two pretests: One to determine how strong our contexts were in terms of speaker's knowledgeability and one to examine the acceptability and coherence of the items to be tested in the eye-tracking experiment. Both pretests were created in Ibex and hosted on Ibex farm (Drummond, 2007). Native speakers of Dutch filled in the online questionnaires voluntarily (N=15 and N=16, respectively). The same forty items were included in both pretests as well as the same thirty-two fillers.

5.3.4.1 Pretest 1

The first pretest aims to establish the strength of our knowledgeability contexts (i.e., speaker ignorance vs. authority contexts) by exploring how compatible those contexts are with a follow-up sentence that provides precise information. Participants were given texts like that in (1) (repeated below, without outro sentence) in all four conditions (Latin square design), with the difference being that the target sentence had the precisifier *precies* ('precisely') in place of *minstens* ('at least').

(4) **Example item**

Intro
Wesley heeft zijn eigen zaak waar hij met veel plezier tatoeages zet. Hij maakt er altijd echte kunstwerken van en het is er dan ook meestal erg druk. Hij wilde deze donderdag daarom hard doorwerken.
'Wesley runs his own tattoo parlor, which he enjoys a lot. He always makes real works of art and it is usually very busy. That's why he wanted to work hard

all through Thursday.'

Context sentence

Speaker ignorance context
Ik weet niet precies hoe het met de drukte zat, maar ik heb wel een idee.
'I don't know exactly how busy it got, but I have got an impression.'

Speaker authority context
Ik weet precies hoe het met de drukte zat en daarom zal ik je erover vertellen.
'I know exactly how busy it got, and I'll tell you about it.'

Target sentence
Wesley $\left\{\begin{matrix}\text{moest}\\ \text{heeft}\end{matrix}\right\}$ *precies zes mensen op donderdag* $\left\{\begin{matrix}\text{tatoeëeren}\\ \text{getatoeëerd}\end{matrix}\right\}$.
Wesley $\left\{\begin{matrix}\text{had to}\\ \text{has}\end{matrix}\right\}$ precisely six people on thursday $\left\{\begin{matrix}\text{tattoo}\\ \text{tattooed}\end{matrix}\right\}$.
'Wesley had to tattoo/tattooed exactly six people on Thursday.'

They were asked to assess how compatible the target sentence was with the preceding context. They did so on a Likert scale from 1 to 7, where 1 stands for "the sentence does not fit the preceding context" and 7 for "the sentence fits well the preceding context". High compatibility scores are taken to indicate that the contexts are strong speaker knowledgeability contexts, whereas low compatibility scores indicate poor speaker knowledgeability contexts; intermediate scores are taken to point to chameleon-like contexts, that is, contexts that are underspecified with respect to speaker's knowledgeability.

Indeed, we found that our speaker ignorance contexts provide a strong cue to speaker ignorance, so they are poor knowledgeability contexts, as they were judged to clash severely with the later use of *precisely* by the speaker (mean: 2.353); on the other hand, authority contexts received significantly higher scores ($p < .0001$), which were also scores at ceiling (mean: 6.077). Apart from this strong context effect, no other effect was found to be significant. Thus, this pretest makes it clear that our context manipulation works and that we have genuine speaker ignorance and speaker authority contexts.

5.3.4.2 Pretest 2

Pretest 2 was designed to test how coherent our test items were, i.e., texts like that in (1). The task was same as in pretest 1. Participants were given texts like (1) in all four conditions (Latin square design) and had to judge how compatible the target sentence with *minstens* ('at least') was with the preceding context. As before, they did so on a Likert scale from 1 to 7, where 1 stands for "the sentence does not fit the preceding context" and 7 for "the sentence fits well the preceding context". The idea was to exclude from the eye-tracking experiment those items that would receive low mean coherence scores overall.

Given our participants' judgements, we excluded eight items that received the lowest mean coherence scores overall (< 4.3).[2] The remaining 32 items (a convenient number for our 2×2 Latin square design), which were judged to consist in coherent texts (their mean score was greater than 4, the middle value of the coherence scale), were included as test items in the eye-tracking experiment.

We also ran an analysis on the overall dataset, without excluding the aforementioned eight items. Interestingly, there was a difference in coherence between items with speaker authority contexts and items with speaker ignorance contexts, with the latter scoring significantly lower than the former (*means* = 4.243 and 5.202, respectively, $p < .001$); no effect of the Verb factor or interaction was attested though. I should point out here that in this experiment no bad/incoherent controls were included; the filler items were all well-formed texts (mean > 5.80) that constituted the test items of another experiment. In the absence of ill-formed, clearly incoherent, control items, e.g., items like the ignorance context items in pretest 1 with *precisely* in the target sentence (*mean* = 2.353; note how higher ignorance contexts score in pretest 2), potential smaller differences between items might have been magnified in pretest 2. This is actually no big surprise. In this type of task, participants' attention is drawn to and focused on the coherence relation between the target sentence and the preceding text, and given that there were no bad/ill-formed control items that participants could compare the (intuitively) well-formed items to, it is very likely that they became (more) fastidious and stricter in their judgements, resulting in the enlargement of possible small differences between the items. It is not clear at the moment what these differences could be due to, but I will come back to them in section 5.4.

Before moving to the presentation of our eye-tracking experiment, we could already evaluate part of the predictions of the theoretical accounts given the results of pretest 2. If anything, we would expect the difference with respect to the speaker Context to be in the opposite direction according to the majority of the accounts of speaker ignorance. On the contrary, the scores the speaker authority contexts received were remarkably high (*mean* = 5.202) and no interaction of the Context and Verb factors was found (see figure 5.1), as was expected according to accounts that derive obligatory speaker ignorance inferences for unembedded uses of *at least*. Specifically, the combinations of a target sentence with a non-modal main verb and an authority context obtained a mean coherence score of 5.110, much higher than the corresponding ignorance context (*mean* = 4.186, see figure 5.1). This is unexpected according to Cohen & Krifka (2014), Geurts & Nouwen (2007), Nouwen (2010), and Spector (2015), which predict a contradiction of the relevant target sentence with a preceding authority context, to be manifested as lower coherence scores like those, for instance, found for ignorance contexts in pretest 1. It is also at odds with Cia-

[2]Mistakingly, we did not exclude an item with overall *mean* = 4.12 and instead we excluded an item with *mean* = 4.43.

rdelli et al.'s (2017), Coppock & Brochhagen's (2013b), and Spychalska's (2015) accounts that likewise predict an inconsistency between target and authority context sentences, but due to hard-to-cancel implicatures associated with the superlative modifier in the target sentence. In contrast, these accounts predict no incompatibility, but rather complete coherence in the relevant speaker ignorance contexts contrary to fact. In sum, pretest 2 results seem to go against what was predicted given all these accounts that derive obligatory speaker ignorance effects and TH1. Moreover, they are in line with accounts that do not impose any restriction on *at least* forcing obligatory, context-independent speaker ignorance (Büring, 2008, Cummins & Katsos, 2010; Kennedy, 2015; Nouwen, 2015; Schwarz, 2016a) and hypothesis TH2, suggesting that no such inference was triggered in the speaker authority Context condition.

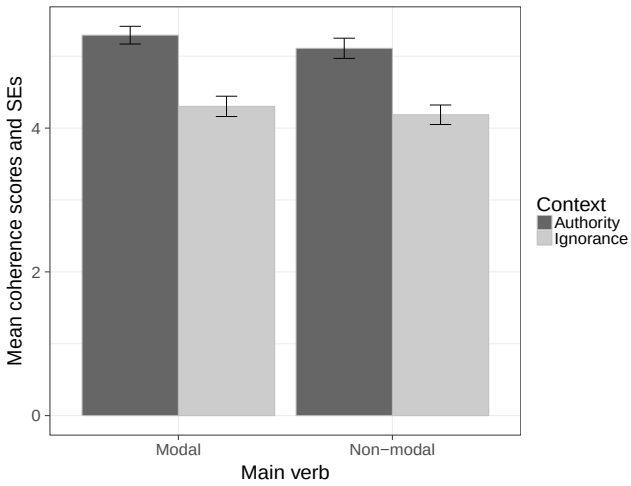

Figure 5.1: Mean compatibility scores of target sentence with preceding context per condition in pretest 2 (eye-tracking experiment 1).

5.3.5 Methods

5.3.5.1 Participants

Forty native speakers of Dutch participated in the eye-tracking experiment (28 female, mean age: 24.23, age range: 18–65), who where recruited from the UiL OTS participant database. Three participants were excluded early on, as two of them turned out to have participated in pretest 2, while the third one was falling asleep during the experiment, resulting in a very limited and messy data record. All participants were paid for their participation (7.5 euro). Lastly, they had normal or corrected to normal vision and were naive as to the purpose of the study.

5.3.5.2 Materials

As already said, we constructed thirty-two different texts as test items, see an example in (1). Those texts consisted of an Intro, the context sentence, the target sentence, and sometimes of a final sentence (Outro), see (1). In fact, in twenty-five of the test items the target sentence was not the last sentence of the text, but was followed by another sentence (Outro). The Intro part introduced the main character of the text and consisted of two or three sentences (extending on three to five lines). The context sentence had the Context manipulation, with two versions of contexts: A context where the speaker had partial knowledge (speaker ignorance context) or a context where the speaker was stating that she was knowledgeable (speaker authority context). The formulation of the context sentences varied across items, but we made sure to keep the contrast between partial knowledgeability vs. full knowledgeability clear and constant across items. Also, we tried to keep the items minimally different with respect to the context sentence in terms of material and length.

Next, the target sentence had the Verb manipulation and the numeral modifier *at least*. As you see in (1), the target sentence starts with a reference to the main character of the text (by name or use of a pronoun), which is followed by the main verb (Verb manipulation: *had to* vs. *has*). Then the modified numeral phrase follows (*at least six people*), which is constant across conditions, and consists of the numeral modifier *minstens* ('at least') and a number phrase. The numbers used in the number phrases across items are described by one to five words, whereas the nouns in these phrases are from one (polysyllabic in most cases) word long to two words long.[3] Both low (e.g., two, seven, fifteen) and higher (e.g., sixty-four, four hundred, hundred thirty million) numbers were used. The numeral modifier *minstens* as well as the number phrase (see section 5.3.6 re splitting into regions of interest) are equally critical as the interpretation of the whole modified numeral phrase is completed after reading the latter phrase. The modified numeral phrase is in turn succeeded by some material (usually a prepositional phrase: e.g., *on Thursday*) also constant across conditions. This is composed of two to six words, which is long enough to serve as a spillover region after the completion of reading the critical modified numeral phrase and before the last word of the sentence, where the conditions diverge again. Specifically, the last verbal material varies depending on the main verb type (i.e., bare infinitive vs. participle) and can potentially trigger sentence wrap-up effects or other effects due to the differences of the relevant material (e.g., with respect to frequency, length, or semantics).

Last, the Outro consists of one sentence, which is the last sentence of the item and the same in all conditions. In Appendix B, you can find a list of all test items.

Our test items appeared in four conditions (given our 2 CONTEXTs × 2 VERBs design). They were rotated through twelve lists in a Latin square design

[3]The number phrases can contain a numeral or a measure phrase. We assume that there is no difference in *minstens* ('at least') modifying a bare numeral or a measure phrase.

and every list had eight items per condition. Every participant saw only one list with 128 trials in total: four practice trials, thirty-two test trials, and ninety-two filler trials. The latter constituted the test items of two separate experiments. The trials of every list were randomly ordered for each participant. Sixty-eight comprehension statements about the story narrated in the texts were included too, to control for whether participants pay attention to the texts they are reading. Participants had to indicate whether the statement was true or false given the text they had just read. These comprehension statements targeted information that was irrelevant to the manipulations of the experiment. Eighteen (out of the 32) experimental items were followed by such a comprehension statement, while the remaining fifty comprehension statements were follow-ups of the filler items.

Experimental items extended on five to eight lines and filler items between one and eight lines. Every line included up to sixty-eight characters (including spaces). The target sentence most of the time was kept on one single line, unless it was more than 68 characters long, in which case a line break occurred after at least two (long enough) words after the numeral modifier *at least*. Context sentences appeared in one or two lines.

5.3.5.3 Procedure

Participants were seated in a comfortable chair in a sound-treated booth and the distance of their head to the screen was 55–70 cm. Although viewing was binocular, for all participants, the position of their right eye was measured and recorded. Specifically, their eye movements were recored with an EyeLink 1000 eye tracker in remote mode (using a target sticker), sampling at 500Hz. The stimuli were presented on a 17-inch Acer AL1717 monitor and a three-button button box was provided to participants for answering the comprehension questions or moving on.

Participants first read the instructions, where they were informed that they would be presented with short stories, each consisting in a short paragraph. After the story, a comprehension question about the text just presented would occasionally appear, which they had to answer by using the button box (right button for YES and left button for NO). Participants had to press the third, middle, button in order to progress. After reading the instructions, the calibration procedure with nine fixation points would start. After a successful calibration, participants would move on to the practice block, where they read four practice items and answered comprehension questions that two of them were associated with. Then the experiment block began. Before the presentation of a stimulus, a fixation point appeared on the screen to mark the beginning of the text and help participants find the start of the first sentence and avoid a seeking behavior that can influence the reading data. Participants were instructed to always fixate on this point, which would disappear once successfully fixated, giving its place to a story. Lastly, participants were instructed to read the stories at a normal pace, as they would do in their everyday life. The whole experiment

lasted approximately 40 minutes.

5.3.6 Results

All 40 participants answered correctly 87% of the comprehension questions on average and no subject scored lower than 75%. The data of all these participants were included in the statistical analyses. Data of two items that contained a typo were excluded from the analyses (.13% of the observations). These two items were not excluded altogether though, as the typos were noticed early on and were corrected for the most part of the experiment. We performed data clean up using the software *Fixation* (Cozijn, 2006). *Fixation* presents the stimuli in the order they appeared in the experiment block, placing the individual fixations over the stimulus. We included all fixations and blinks for statistical analysis, unless they were clearly outside the stimulus area and/or disrupted the reading pattern in an illogical way. In addition, another .21% of the total observations was excluded because of problematic trials (trials with few or no fixations due to track losses or participants' sloppy reading). For all measures, fixations of 0 ms were excluded from the analysis. We log-transformed the remaining reading time data to approximate normality.

We split the texts of the test items into regions for the purpose of the analyses. The intro constituted Region 1 and the speaker context was Region 2. The target sentence was broken down into smaller regions as illustrated in Table 5.4 for (1). The numeral was grouped together with the noun (Region 6) to comply with the syntactic structure Krifka (1999) and Geurts & Nouwen (2007) propose for modified numeral phrases. The prepositional phrase following Region 6 was long enough to serve as a spillover region (see section 5.3.5.2), given that Region 6 is where the interpretation of the whole modified numeral phrase is completed. The outro sentence of the texts was not included for analysis, because not all items contained such a part.

Region 3	**Region 4**	**Region 5**	**Region 6**	**Region 7**	**Region 8**
Wesley	{had to / has}	at least	six people	on Thursday	{tattoo / tattooed}.

Table 5.4: Eye-tracking experiment: Regions of target sentence.

For each of those regions we analyzed the following seven reading time measures: (i) *first-pass time*, which is the sum of fixations in a region before exited for the first time to any direction, (ii) *right-bounded time*, which is the sum of fixations in a region before exited for the first time to the right, (iii) *regression path duration*, i.e., the sum of fixations since the first fixation in a region until it is exited to the right for the first time, (iv) *probability of regression*, i.e., the percentage of regressive eye-movements out of a region after a first-pass fixation in a region (v) *total reading time*, which is the sum of all fixation durations in a region, including 2nd, 3rd, ..., nth pass, (vi) *re-reading time*, that is, the sum

of fixations in a region excluding first-pass fixations, and (vii) *probability of re-reading*, i.e., the percentage of entering a region for a 2nd, 3rd, ..., nth time. First-pass, probability of regression, and right-bounded times are taken to be early measures, re-reading time and probability are assumed to be late measures, while regression path duration and total reading time are intermediate measures (they may reflect both first and second pass time). Raw means of the eye-tracking measures per region are reported in tables 5.5 and 5.6. Note that we report on values for regions that were fixated after reading later, critical, parts of the individual text that include the Context and Verb manipulations.

We conducted linear mixed-effects regression analyses on the log-transformed reading-time data and mixed-effects logistic regression analyses on the categorical regression probability and re-reading probability measures using the `lme4` package (Bates, Mächler, Bolker & Walker, 2015) in `R`. All analyses included two predictors, speaker Context (authority vs. ignorance) and Verb (modal vs. non-modal), which were treatment-coded, with authority and non-modal as the reference levels, respectively. The analyses also included random intercepts and slopes for participants and items for all fixed effects. We applied backward model selection for random effects (Barr et al., 2013) and in the following I will be presenting the output of the model with the maximal random-effect structure that converged and had the best fit. Finally, apart from determining whether the speaker knowledgeability Context affects the incremental interpretation of (unembedded and embedded) *at least*-modified numerals, we also aimed to determine whether it does so differentially for the two Verb conditions. Therefore, the interaction of the two factors was generally kept in the models, even if it was not significant but had the expected sign (see Gelman & Hill, 2007, p. 69), cf. the predictions in section 5.3.3. *P*-values are calculated based on Satterthwaite's approximations, using the *ImerTest* R package.

5.3.6.1 Region 1, Intro

The Verb factor was found to have a significant effect on re-reading probability and a marginal one on re-reading time such that the Intro was more likely to be re-read and tended to be re-read for longer when the main verb of the target sentence was modal than when it was non-modal (re-reading probability: $\beta = .487$, $SE = .213$, $z = 2.287$, $p < .05$, re-reading time: $\beta = .251$, $SE = .122$, $t = 2.049$, $p = .056$).

Moreover, there was a significant positive effect of Context on the probability of re-reading ($\beta = .525$, $SE = .217$, $z = 2.420$, $p < .05$) and on re-reading times ($\beta = .238$, $SE = .109$, $t = 2.180$, $p < .05$) in the present region as well as a significant negative interaction between the Verb and Context factors in these two measures (re-reading probability: $\beta = -.627$, $SE = .291$, $z = -2.157$, $p < .05$, re-reading time: $\beta = -.552$, $SE = .177$, $t = -3.125$, $p < .01$). That is, there is a significant difference between ignorance and authority Context conditions, with the former having a greater inhibitory effect, and this difference is smaller in the modal condition than in the non-modal Verb condition.

Condition		Region			
Context	Verb	Region 1 Intro	Region 2 Context	Region 3 *Wesley*	Region 4 Verb
First pass (ms)					
Authority	non-modal			370	266
Authority	modal			378	267
Ignorance	non-modal			426	248
Ignorance	modal			384	276
Right bounded (ms)					
Authority	non-modal			376	286
Authority	modal			406	290
Ignorance	non-modal			440	266
Ignorance	modal			404	310
Regression path duration (ms)					
Authority	non-modal			387	384
Authority	modal			486	395
Ignorance	non-modal			466	370
Ignorance	modal			451	423
Probability of regression					
Authority	non-modal		.885	.018	.313
Authority	modal		.882	.034	.302
Ignorance	non-modal		.874	.028	.280
Ignorance	modal		.889	.075	.269
Total reading time (ms)					
Authority	non-modal	6978		376	313
Authority	modal	7355		428	359
Ignorance	non-modal	7253		434	311
Ignorance	modal	7151		424	396
Re-reading time (ms)					
Authority	non-modal	1207		317	260
Authority	modal	1651		393	343
Ignorance	non-modal	1667		400	310
Ignorance	modal	1267		387	411
Probability of re-reading					
Authority	non-modal	.300	.240	.478	.395
Authority	modal	.368	.232	.508	.482
Ignorance	non-modal	.379	.246	.460	.380
Ignorance	modal	.355	.260	.531	.453

Table 5.5: Means of eye-tracking measures for Regions 1, 2, 3, and 4 in different conditions in eye-tracking experiment 1.

Condition		Region			
Context	Verb	Region 5 *at least*	Region 6 *six people*	Region 7 PP	Region 8 PTCP/INF
First pass (ms)					
Authority	Non-modal	236	549	388	295
Authority	Modal	248	518	402	285
Ignorance	Non-modal	228	548	379	292
Ignorance	Modal	239	576	398	275
Right bounded (ms)					
Authority	Non-modal	263	605	433	336
Authority	Modal	298	632	447	321
Ignorance	Non-modal	615	584	431	340
Ignorance	Modal	290	681	449	326
Regression path duration (ms)					
Authority	Non-modal	341	682	568	708
Authority	Modal	432	755	733	946
Ignorance	Non-modal	388	679	566	808
Ignorance	Modal	421	808	649	793
Probability of regression					
Authority	Non-modal	.195	.134	.171	.386
Authority	Modal	.274	.218	.215	.421
Ignorance	Non-modal	.242	.139	.183	.360
Ignorance	Modal	.257	.217	.213	.391
Total reading time (ms)					
Authority	Non-modal	317	730	522	339
Authority	Modal	379	778	557	327
Ignorance	Non-modal	323	746	534	363
Ignorance	Modal	399	838	581	338
Re-reading time (ms)					
Authority	Non-modal	289	587	427	303
Authority	Modal	356	629	436	370
Ignorance	Non-modal	318	537	473	332
Ignorance	Modal	398	626	505	365
Re-reading probability					
Authority	Non-modal	.359	.375	.380	.182
Authority	Modal	.453	.473	.424	.205
Ignorance	Non-modal	.361	.444	.383	.239
Ignorance	Modal	.466	.482	.410	.206

Table 5.6: Means of eye-tracking measures for Regions 5, 6, 7, and 8 in different conditions in eye-tracking experiment 1.

5.3.6.2 Region 2, Speaker knowledgeability context

For this region, basically, we only analyzed the data for the probability of re-reading measure, which reflects how likely it is for the readers to go back to the context sentence after having read later parts of the text (possibly including our manipulations) as well as for the probability of regression measure, which indicates how likely it is for the readers to regress to the Intro part of the text. Looking into the reading time measures would not be any informative, as the two types of speaker Context differ (to some extent) as to the lexical material used in that region or sometimes the sentence length, which might influence and confound reading times. No factor turned out to have an effect on the probability of re-reading or the probability of regression.

5.3.6.3 Region 3, *Wesley*

There was no effect on any measure in Region 3, which is the first region of the target sentence, only consisting of the subject.

5.3.6.4 Region 4, Main verb

No significant effect was detected in Region 4 containing the Verb manipulation (modal vs. non-modal), except for a positive main effect of Verb in re-reading probability that was close to being significant (β = .390, SE = .202, z = 1.932, p = .053), and that was marginal in re-reading (β = .156, SE = .089, t = 1.747, p = .087) and total reading times (β = .093, SE = .051, t = 1.835, p = .071). That is, readers tended to re-read the main verb of the target sentence more often and for longer when this was a modal rather than a non-modal verb and to spend overall more time reading the main verb in the former case.

5.3.6.5 Region 5, *at least*

A positive effect of Verb was revealed in the region of *at least* right after the verb region in various measures. Participants would slow down when first reading *at least* before reading further (right-bounded: β = .106, SE = .040, t = 2.689, p < .01), and including going back to previous regions (regression path duration: β = .145, SE = .068, t = 2.144, p < .05), and were more likely to regress to previous regions (β = .535, SE = .237, z = 2.262, p < .05) when the verb in the region they had just read was modal rather than non-modal. The same inhibitory effect due to the presence of a modal verb was exhibited on total reading times (β = .148, SE = .045, t = 3.246, p < .01), and on re-reading times (β = .191, SE = .084, t = 2.271, p < .05) and probability (β = .470, SE = .191, t = 2.461, p < .05).

5.3.6.6 Region 6, *six people*

The disruptive effect of the modal condition continued showing up in Region 6; specifically, in the probability of regression (β = .688, SE = .231, z = 2.978, p < .01) and of re-reading (β = .486, SE = .180, z = 2.695, p < .01) measures. Participants were more likely to regress to previous regions once reading *six people* for the first time and also more likely to re-read that region when the target sentence had a modal as opposed to a non-modal verb (two regions earlier). In addition, a positive effect of the Context condition was observed in the present region in the probability of re-reading (β = .366, SE = .183, z = 1.994, p < .05): *six people* was re-read more frequently when the target sentence followed a speaker ignorance context than a speaker authority context (baseline Verb condition: non-modal, see figure 5.2). No other effect was found to be significant.

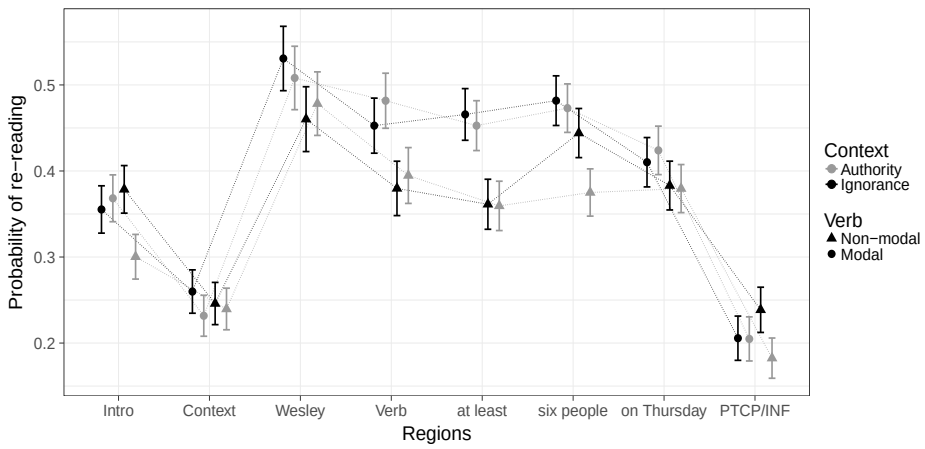

Figure 5.2: Probability of re-reading each region per condition.

5.3.6.7 Region 7, PP

The analyses revealed no effect on any of the measures under investigation in the present region, except for a marginal positive effect of Verb in regression probability (β = .401, SE = .236, z = 1.701, p = .089).

5.3.6.8 Region 8, PTCP/INF

There were no main or interaction effects in the last region.

5.4 Discussion

Effects of Modal. Let us first start with the hindering effect of the modal verb observed during the incremental interpretation of test items. This effect first occurs at the region of *at least* (Region 5) right after the region with the Verb manipulation and appears to disrupt (early and late) reading until Region 7 (PP). It also shows up in late measures of previously read regions like Region 4 (verb) and Region 1 (Intro). That is, after encountering a modal verb in Region 4 and while reading the subsequent Regions 5, 6, and 7 for the first time readers are likely to regress back to and re-read previously read material. This lasting disruptive effect does not seem to relate in any way to the purposes of the present study though. Rather, it is most possibly caused due to a number of aspects that make the modal verbs used in the modal Verb condition (*moest* 'had to'/*wilde* 'wanted to') different and perhaps more complex compared to the auxiliary verbs used in the non-modal Verb condition (*heeft* 'has'/*is* 'is'), e.g., word frequency, word length, tense, semantics, etc.

Interaction effect. According to our results, it does not make any difference whether the target sentence following an ignorance context has an embedded or unembedded occurrence of *at least*. Hence, our findings do not confirm hypothesis PH10 (see Table 5.7 repeating Table 5.3 for convenience) and we could firmly say that we found no evidence indicating that an additional cost is associated with the interpretation of *at least* in embedding environments in the ignorance condition due to the extra operation of covert wide-scope movement of *at least* relative to the present universal modal operator.

Hypotheses		**Predictions re embedded *at least***
PH10	Wide-scope quantifier movement is costly	Processing penalty at *at least* phrase in Modal condition of **ignorance contexts**
PH11	Wide-scope quantifier movement is not a costly process	No interaction effect in **target sentence** due to movement

Table 5.7: Processing hypotheses and predictions re embedded occurrences of *at least*.

Our findings are then in line with hypothesis PH11, assuming that we follow the majority of the literature, which requires a wide-scope configuration for the derivation of an ignorance interpretation with an embedded use of a superlative modifier. Of course, we cannot exclude the possibility that a more sensitive method than that used here might detect on-line traces of wide-scope movement of *at least* in embedding contexts.

Nevertheless, we did find an interaction effect in Region 1 (Intro) and this

is the only interaction effect attested in the present experiment. Such an effect was manifested in re-reading probability as well as in re-reading time, and it was negative: Readers re-read the Intro part of the test items more frequently and for longer after encountering a non-modal verb in a target sentence that followed an ignorance context, that is, the difference between ignorance and authority conditions in these two measures was bigger in the non-modal condition than in the modal condition.

Importantly, it was also found that readers were likely to go back to the Intro after having read an ignorance context as opposed to an authority one (main effect of ignorance Context in re-reading probability). One explanation for such an effect could be that there might be an authority or knowledgeability bias set up in the Intro, which is revisited or questioned after having read the ignorance context sentence. That is, the author of the texts makes some informative introduction at the beginning of each text, when she presents the main character and an activity/situation that character is involved in, creating the impression of a fully and well-informed narrator. So it might be the case that participants having read the ignorance context sentence return to the Intro to reconsider that impression; this does not happen in the authority contexts because those are consistent with the bias in question.

Now, going back to the interaction effects, if we assume that such a bias is indeed created by the Intro, the attested negative interactions could be interpreted as follows: The fact that readers are even more likely to revisit the information in the Intro in the non-modal condition and reconsider it for longer could suggest that the speaker ignorance signal triggered in such a condition is more robust than in the modal condition, prompting more strongly for reconsideration of the knowledgeability bias. But why would the speaker ignorance signal be stronger in the non-modal condition, given that the ignorance context gives the same prompt to derive an ignorance implicature in the two Verb conditions? Why would participants derive ignorance implicatures more robustly in the non-modal than in the modal Verb condition?

Ignorance contexts should prompt for a wide-scope interpretation of the target sentence in the modal Verb condition. However, this configuration of the interaction of a superlative modifier with a universal modal is less preferred compared to the more frequent and salient narrow-scope configuration. Given that the truth-conditional content in the latter case (*in every deontically accessible world, the number n of people Wesley tattooed is* ≥ 6, e.g., for (1)) is compatible with the information expressed by the ignorance contexts, it seems that some participants prefer to stick to the more frequent narrow-scope configuration, which is not the right configuration for deriving ignorance interpretations (according to the majority of the accounts). As a result, less robust ignorance implicatures are derived in the modal condition of ignorance items, and the difference from the authority items becomes smaller. That is, this difference is smaller as compared to the non-modal condition, where the need for revisiting the Intro with the knowledgeability bias is greater due to the more robust ignorance profile of the speaker created after the consistent

derivation of an ignorance implicature in the relevant target sentence.

Interestingly, this difference in robustness of speaker ignorance between the two Verb conditions was not attested in any of the critical regions.

Context effect. The effect that is primarily related to our manipulations and purposes of the present study is that occurring at Region 6, the critical region comprising the number and noun of the modified numeral phrase, where the interpretation of the whole phrase is completed. Readers were more likely to re-read the region of *six people* when the target sentence followed an ignorance context sentence than an authority one. Apparently, there is something troublesome computationally about this region in combination with the ignorance context that makes participants revisit that region. What could have caused this effect? In the following, we will consider in which cases the attested effect of ignorance Context should arise evaluating various (combinations of) hypotheses posited in sections 5.3.2 and 5.3.3 given different theoretical proposals. The different options to be considered will be presented schematically and will be assessed individually.

Option 1

Theoretical assumption: Speaker ignorance is derived as a semantic implication (Geurts & Nouwen, 2007; Nouwen, 2010; Spector, 2015)

Processing assumptions: PH1: Rapid derivation of semantic inferences
PH6: Semantic violation/contradiction is taxing

Given **option 1**, the opposite Context effect of what we found should be attested (i.e., negative effect of Context), due to the processing cost incurred by the contradiction of the target sentence with the authority context sentence.

Option 2

Theoretical assumption: Speaker ignorance is derived as a semantic implication (Geurts & Nouwen, 2007; Nouwen, 2010; Spector, 2015)

Processing assumptions: PH2: Costly derivation of semantic inferences
PH6: Semantic violation/contradiction is taxing

Similarly to **option 1**, the opposite Context effect of the attested one should be observed (i.e., negative effect of Context). That is, an *extra* processing cost should be associated with the authority Context condition due to contradiction.

Hypotheses	**Predictions** re unembedded *at least*
PH1 Rapid derivation of	No precessing penalty
PH2 Costly derivation of	Processing penalty at *at least* phrase
semantic implications:	in **ignorance/authority contexts**
• due to lexical semantics	(GN)
• due to last-resort insertion of covert existential modal	(N10)
• due to obligatory LF exhaustification operation	(Spec)
PH3 Rapid derivation of	No precessing penalty
PH4 Costly derivation of	Processing penalty at *at least* phrase
pragmatic implications:	in **ignorance contexts**
• via Quality maxim	(CB, CCR)
• via assertibility conditions	(Sp)
• via Quantity maxim	(B, CKr, K, N15, S)
PH5 *At least* is psychologically complex due to disjunctive status	Processing penalty at *at least* phrase in **ignorance/authority contexts** (CK)
PH6 Semantic violation/contradiction is taxing	Processing penalty at *at least* phrase in **authority contexts**
PH7 Semantic violation/contradiction is not taxing	No processing penalty
	(GN, N10, Spec, CKr)
PH8 Inconsistency due to hard-to-cancel implicature is taxing	Processing penalty at *at least* phrase in **authority contexts**
PH9 Inconsistency due to hard-to-cancel implicature is not taxing	No processing penalty
	(CB, CCR, Sp)

Table 5.8: Processing hypotheses and predictions re unembedded occurrences of *at least*. B stands for Büring (2008), CCR for Ciardelli et al. (2017), CKr for Cohen & Krifka (2014), CB for Coppock & Brochhagen (2013b), CK for Cummins & Katsos (2010), GN for Geurts & Nouwen (2007), K for Kennedy (2015), N10 for Nouwen (2010), N15 for Nouwen (2015), S for Schwarz (2016a), Spec for Spector (2015), and Sp for Spychalska (2015).

Option 3

Theoretical assumption: Speaker ignorance is derived as a semantic implication (Geurts & Nouwen, 2007; Nouwen, 2010; Spector, 2015)

Processing assumptions: PH1/2: Rapid/Costly derivation of semantic inferences
PH7: Semantic violation/contradiction is not taxing

In **option 3**, although PH7 is tremendously unlikely given a large number of experimental findings (see, e.g., Albrecht & O'Brien, 1993; Baker & Anderson, 1982; Bohan & Sanford, 2008; Myers, O'Brien, Albrecht & Mason, 1994; van Berkum et al., 1999, 2003), no difference is expected between the two Context conditions, contrary to fact.

Consequently, our finding does not confirm any of the processing hypotheses formulated for a semantic analysis of speaker ignorance, hence it does not support a semantic account of speaker ignorance, such as Geurts & Nouwen's (2007) and Nouwen's (2010) accounts, and Spector's (2015) grammatical account of ignorance. Let us know continue with the the rest of the possible options.

Option 4

Theoretical assumption: Superlative modifiers are disjunctive (Cummins & Katsos, 2010)

Processing assumptions: PH5: *At least* is psychologically complex due to its disjunctive status (Cummins & Katsos, 2010)

The observed positive effect of Context goes, moreover, against **option 4**, because no difference between the two types of Context would be expected. If we additionally consider hypothesis PH3, as in **option 5** below, the observed effect is again left unexplained (for convenience, I am repeating the table with the theoretical hypotheses and predictions in Table 5.9 on the next page).

Option 5

Theoretical assumptions: Superlative modifiers are disjunctive (Cummins & Katsos, 2010)
TH2: Ignorance inferences are non-obligatory

Processing assumptions: PH5: *At least* is psychologically complex due to its disjunctive status (Cummins & Katsos, 2010)
PH3: Pragmatic inferences are rapidly computed

Hypotheses	**Predictions re unembedded *at least***
TH1 Obligatory ignorance	**Target** compatible with **Ignorance context** **Target** incompatible with **Authority context** due to semantic/pragmatic violation (CCR, CKR, CB, GN, N10, SPEC, SP)
TH2 Non-obligatory ignorance	**Target** compatible with **Ignorance context** **Target** compatible with **Authority context** (B, CK, K, N15, S)

Table 5.9: Hypotheses re obligatoriness of ignorance implication and predictions given the design of Experiment 1. B stands for Büring (2008), CCR for Ciardelli et al. (2017), CKR for Cohen & Krifka (2014), CB for Coppock & Brochhagen (2013b), CK for Cummins & Katsos (2010), GN for Geurts & Nouwen (2007), K for Kennedy (2015), N10 for Nouwen (2010), N15 for Nouwen (2015), S for Schwarz (2016a), SPEC for Spector (2015), and SP for Spychalska (2015).

Our effect can instead be explained by **option 6**, **option 7**, or **option 8**, that is, either we assume that superlative modifiers have a disjunctive status (Büring, 2008) and this makes them psychologically more complex (Cummins & Katsos, 2010), or not (Kennedy, 2015; Nouwen, 2015; Schwarz, 2016a).

Option 6

Theoretical assumptions: superlative modifiers are disjunctive (Cummins & Katsos, 2010)

TH2: Ignorance inferences are non-obligatory

Processing assumptions: PH5: *At least* is psychologically complex due to its disjunctive status (Cummins & Katsos, 2010)

PH4: Pragmatic inferences are costly

Option 7

Theoretical assumptions: superlative modifiers are disjunctive (Büring, 2008)

TH2: Ignorance inferences are non-obligatory

Processing assumptions: PH4: Pragmatic inferences are costly

Option 8

Theoretical assumptions: TH2: Ignorance inferences are non-obligatory (Kennedy, 2015; Nouwen, 2015; Schwarz, 2016a)

Processing assumptions: PH4: Pragmatic inferences are costly

Let us now turn to the options where ignorance inferences are obligatory and context-independent implicatures as predicted by a number of pragmatic proposals.

Option 9

Theoretical assumption: TH1: Ignorance inferences are obligatory (Ciardelli et al., 2017; Coppock & Brochhagen, 2013b; Spychalska, 2015)

Processing assumptions: PH3: Rapid derivation of pragmatic inferences
PH8: Inconsistency due to hard-to-cancel implicature is taxing

Option 10

Theoretical assumption: TH1: Ignorance inferences are obligatory (Ciardelli et al., 2017; Coppock & Brochhagen, 2013b; Spychalska, 2015)

Processing assumptions: PH3: Rapid derivation of pragmatic inferences
PH9: Inconsistency due to hard-to-cancel implicature is not taxing

According to **option 9**, which includes the theoretical assumptions made by Ciardelli et al. (2017), Coppock & Brochhagen (2013b), and Spychalska (2015), the opposite Context effect should be attested, due to the cost associated with the relevant resulting inconsistency in the authority condition, while no difference is expected between the ignorance and authority Context conditions, given **option 10**, where we take PH9 instead of PH8. What happens if the derivation of ignorance inferences is computationally demanding apart from obligatory?

Option 11

Theoretical assumption: TH1: Ignorance inferences are obligatory (Ciardelli et al., 2017; Coppock & Brochhagen, 2013b; Spychalska, 2015)

Processing assumptions: PH4: Pragmatic inferences are costly
PH8: Inconsistency due to hard-to-cancel implicature is taxing

Given **option 11**, as the cost due to the inconsistency in question is an extra cost to that caused by the effortful derivation of the ignorance inference in the Authority condition, the opposite Context effect of what we found should have been attested. However, keeping PH4 and replacing PH8 with PH9, as in **option 12**, it seems that we arrive at a combination of hypotheses that can explain the observed positive Context effect.

Option 12

Theoretical assumption: TH1: Ignorance inferences are obligatory (Ciardelli et al., 2017; Coppock & Brochhagen, 2013b; Spychalska, 2015)

Processing assumptions: PH4: Pragmatic inferences are costly
PH9: Inconsistency due to hard-to-cancel implicature is not taxing

Hence, under **option 12**, our on-line finding is compatible with the theoretical pragmatic accounts that predict an inconsistency due to hard-to-cancel implicature in the authority Context condition. We arrive at the same conclusions if we add hypothesis PH5 (*at least is psychologically complex due to its disjunctive status*) to each of the above options including the theoretical assumption that the derivation of ignorance is obligatory as in Ciardelli et al. (2017) and Coppock & Brochhagen (2013b), who particularly assume a disjunctive status for superlative modifiers. Crucially, we should be careful with the interpretation of our on-line finding, as the results of pretest 2 reveal that *no* such inconsistency takes place in the authority condition. Recall, that the authority condition received high scores of coherence rates (*mean* = 5.202), allowing us to reject TH1, and, thus, also the theoretical accounts under consideration. This furthermore explains why **option 12** is not tenable in the end.

The next options to be considered concern Cohen & Krifka's (2014) account.

Option 13

Theoretical assumption: TH1: Ignorance inferences are obligatory (Cohen & Krifka, 2014)

Processing assumptions: PH3: Pragmatic inferences are rapid
PH6: Semantic contradiction is taxing

Option 14

Theoretical assumption: TH1: Ignorance inferences are obligatory (Cohen & Krifka, 2014)

Processing assumptions: PH4: Pragmatic inferences are costly
PH6: Semantic contradiction is taxing

Option 15

Theoretical assumption: TH1: Ignorance inferences are obligatory (Cohen & Krifka, 2014)

Processing assumptions: PH4: Pragmatic inferences are costly
PH7: Semantic contradiction is not taxing

The observed ignorance Context effect cannot be explained by **option 13**; under this option, the opposite effect should arise. The observed effect is not compatible with **option 14** either, because this option, too, would predict that a greater processing penalty should be attested in authority contexts. Lastly, contrary to the observed effect, **option 15**, which includes the far-fetched hypothesis PH7 in place of PH6, would predict no difference to arise between ignorance and authority contexts. As it turns out, the Context effect we found cannot be accommodated by Cohen & Krifka's (2014) account plus any additional processing assumptions.

Option 7 and **option 8**, discussed above, basically pertained to Büring's (2008), Kennedy's (2015), Nouwen's (2015), and Schwarz's (2016a) theoretical accounts. We saw that the observed Context effect follows naturally from these two options. If we replace hypothesis PH4, in these options, with hypothesis PH3, we get **options 16** and **17**, respectively, the last ones to be considered:

Option 16

Theoretical assumptions: Superlative modifiers are disjunctive (Büring, 2008)
TH2: Ignorance inferences are non-obligatory

Processing assumptions: PH3: Pragmatic inferences are rapid

Option 17

Theoretical assumptions: TH2: Ignorance inferences are non-obligatory (Kennedy, 2015; Nouwen, 2015; Schwarz, 2016a)

Processing assumptions: PH3: Pragmatic inferences are rapid

As is obvious, the attested Context effect cannot be explained under the last two options, since no Context effect would be predicted to arise.

After examining the various options, we conclude that the observed effect of Context is compatible with **option 6**, **option 7**, **option 8**, and **option 12**. What all these options have in common is hypothesis PH4. Consequently, the processing penalty attested in the ignorance Context condition reflecting that an effortful process takes place at Region 6 (*six people*) is to be associated with the costly computation of the ignorance implicature that the particular context prompts for. More specifically, as already discussed in Section 5.3.1, it is to be associated with the derivation of the specific speaker ignorance implicature, as the context (in combination with the compositional meaning of the target sentence) already entails underspecified speaker ignorance.

More importantly, taking together the results of pretest 2 and those of the main experiment, we further exclude **option 12** from the above options, as already highlighted. This is so, because the results of pretest 2 do not evince any inconsistency between the target and the context sentences in the authority condition, hence hypothesis PH9 (*inconsistency due to hard-to-cancel implicature is not taxing*) is not supported. Therefore, we are left with three options (6, 7, and 8) that concern the neo-Gricean Quantity-based accounts of ignorance implications (Büring, 2008; Cummins & Katsos, 2010; Kennedy, 2015; Nouwen, 2015; Schwarz, 2016a), whereby ignorance is a non-obligatory, context-dependent inference.

To conclude, as has become clear, our overall findings invalidate all theoretical accounts that hypothesize an obligatory derivation of speaker ignorance inferences and thereby predict some (semantic or pragmatic) incompatibility of the target sentence of the test items with a preceding strong authority context.[4] These affected proposals are Ciardelli et al.'s (2017), Cohen & Krifka's (2014), Coppock & Brochhagen's (2013b), Geurts & Nouwen's (2007), Nouwen's (2010), Spector's (2015), and Spychalska's (2015) accounts. As such, our results are in line with Cremers et al.'s (p.c.) very fresh experimental findings revealing that *at least* can (highly) appropriately be used to refer to a certain quantity in scenarios where the speaker is fully knowledgeable of the precise quantity, casting doubt on the supposed obligatory nature of ignorance inferences. Finally, our findings are in favor of neo-Gricean accounts such as Büring (2008); Cummins & Katsos (2010); Kennedy (2015); Nouwen (2015); Schwarz (2016a), whose pragmatic mechanism is not put to use with every (unembedded) occurrence of a superlative modifier.

[4]Recall that pretest 1 demonstrated that the constructed authority contexts were strong speaker knowledgeability contexts, while the ignorance contexts were poor knowledgeability contexts.

Possible confounds and alternative explanations. Before ending this discussion, it is worth pondering a bit more over the ignorance effect attested in the critical Region 6, which is the most interesting effect with respect to the goals of the present study. Although we take it to be an indication of the costly on-line derivation of the specific speaker ignorance implicature that neo-Gricean accounts predict, we cannot disregard a possible alternative explanation of the relevant processing penalty. As already mentioned in section 5.3.5.2, we used both low (e.g., two, seven, fifteen) and higher (e.g., sixty-four, four hundred or hundred thirty million) numbers in our target sentences, but importantly we did not systematically control for the roundness or the preciseness of the numbers used. Let me explain what these terms are taken to refer to. By *roundness* I mean how round the number itself is, for instance, intuitively, 100 is rounder than 103. Round numbers are claimed to be associated with scales with coarser granularity as opposed to non-round numbers (see Cummins, Sauerland & Solt, 2012, a.o.). By the term *preciseness* I refer to how round/imprecise a number is when taking also into account the noun it combines with and specifically the granularity level it sets. For example, 45 *minutes* is rounder/less precise than 40 *minutes* because it is an established and prominent (conceptual) time unit (see Krifka, 2009), which is obviously not the case for 45 *people* vs. 40 *people*.

Given these notions, one could expect that the use of *at least* would be incompatible or less acceptable with a non-round or precise number in speaker ignorance contexts, as it would perhaps be weird for an ignorant speaker (with partial knowledge of the number at stake) to utter such a number with *at least*. Hence, the effect in Region 6 could be because participants realize during the incremental interpretation of the target sentence after reading Region 6 for the first time that the use of the particular material in that region is unnatural and odd given an ignorant speaker that has made an *at least*-utterance. I label this confound as CONF1. Note that this possibility could also perhaps explain the finding of pretest 2: Because of the unexpected and unnatural use of a non-round or precise number with *at least* by an ignorant speaker, participants found the target sentence less compatible with the preceding context when this was a speaker ignorance rather than a speaker authority context.

There are still some further potential interpretations of the effect we find in the critical region, which occurred to me once I considered a couple of intuitive comments I received by Dutch informants that checked the test items of experiment 1. Some have commented that it might still be somewhat unexpected that the ignorant speaker mentions a number in the target sentence (regardless of roundness), despite the fact that she declares that she does not have precise knowledge. On this assumption, the observed Context effect indicates that the readers puzzle over the use of the quantity expression in the target sentence, which makes them revisit the region *six people*. I will dub this confound CONF2. One could possibly expect to find such an (unexpectedness) effect in early measures as opposed to the late measure of re-reading probability, namely, such an effect should perhaps occur rapidly.

Interestingly and relatedly, another comment points out that the use of a

numeral phrase modified by *minstens* gives quite specific and detailed information given the underspecified ignorance of the speaker of ignorance contexts. It was further specified that a numeral modifier like *approximately* instead would be more preferred given the particular ignorance contexts. Given that, the attested Context effect could then be taken to reflect that readers return to the numeral phrase because on second thought they find the information it expresses too specific given that expressed by the preceding ignorance context. I will call this confound CONF3. Table 5.10 summarizes the three confounds discussed.

Confounds

CONF1	No control for preciseness/roundness of number modified by *minstens*
CONF2	Unexpectedness of use of a number in the target sentence
CONF3	Information of *minstens*-phrase too specific given underspecified ignorance from context

Table 5.10: Potential confounding factors of ignorance Context effect in Region 6 in experiment 1.

One more time, if one were to connect the results of pretest 2 and those of the eye-tracking experiment—although the former is a different type of task lacking incoherent/ill-formed control items and does not exhibit the pervasive disruptive effect of modal Verb that the main experiment reveals—both CONF2 and CONF3 could be taken to explain the lower coherence rates the ignorance Context condition received relative to the authority Context condition in pretest 2.

To conclude, the current chapter aimed to answer the following questions, posited in the Introduction (chapter 1):

- ▷ How are ignorance inferences accessed in incremental interpretation of modified numerals? What are the insights into the nature and strength of speaker ignorance inferences and into the underlying mechanism of derivation?

Our overall findings suggest that ignorance effects associated with the superlative numeral modifier *at least* are non-obligatory, context-dependent implications that belong to the specific ignorance type, derived via a pragmatic mechanism like that lying in the core of the Quantity-based Gricean proposals of

ignorance. Our findings are further taken to suggest that their derivation is accompanied by a processing cost, attested in late measures in the region where the whole modified numeral interpretation is completed. I also considered possible alternative explanations of our main finding raising a number of issues. In order to resolve and improve on these issues and make our conclusion firmer as far as the main finding of the present study is concerned, we carried out a follow up eye-tracking reading experiment (experiment 2). The test items are modified so as to only have round numbers in the target sentence that are made relevant by the introduction of (a comparison with) another number at the beginning of the text (Intro). Moreover, in experiment 2 we consider thoroughly the last potential confound (CONF3) and particularly focus on what is perceived as being specific about the meaning of an *at least* modified numeral given the particular ignorance contexts, modifying the experimental design accordingly. Experiment 2 will be presented in chapter 6 that follows.

CHAPTER 6

The time course of speaker ignorance effects: Eye-tracking experiment 2

6.1 Introduction

The experiment presented in this chapter primarily aims to clarify and purify the positive effect of the Context factor attested in the previous experiment (chapter 5): i.e., the processing penalty exhibited at the numeral phrase (or number+noun region) following *minstens* ('at least') in contexts with an ignorant as opposed to an authoritative speaker. More precisely, it does away with CONF1 (see table 6.1), i.e., the confound of the roundness/preciseness of the number being modified in the target sentence as well as with the possibility that its unanticipated mention by an (partially) ignorant speaker (CONF2) interfered with the incremental interpretation of the target sentence. Crucially, the present study zooms in on the interesting remark that the numeral phrases modified by *minstens* ('at least') convey rather specific information, which is perceived as too specific given the (type of) information expressed by the ignorance context sentences of experiment 1 (CONF3).

In the following, I present experiment 2, starting with the modifications of the items of experiment 1 implemented in order to deal with the potential confounding factors CONF1, CONF2, and CONF3, and continue with the updated design of the experiment (section 6.2). In sections 6.3 and 6.4, I discuss the predictions of experiment 2 and report on the acceptability tests performed on the new test items, respectively. Section 6.5 details the methods of the present study, and sections 6.6 and 6.7 present its results and a thorough discussion thereof, respectively. Section 6.8 concludes the present chapter as well as our

study of speaker ignorance implications of numeral modifiers.

Confounds	
CONF1	No control for preciseness/roundness of number modified by *minstens*
CONF2	Unexpectedness of use of a number in the target sentence
CONF3	Information of *minstens*-phrase too specific given underspecified ignorance from context

Table 6.1: Potential confounding factors of ignorance Context effect in Region 6 in eye-tracking experiment 1.

6.2 Ignorance experiment 2: Design

Before presenting the updated design of the present experiment, I would first like to motivate and discuss some changes I made to the design and setup of experiment 1 taking into account the confounding factors listed on table 6.1 above. I will start with CONF3 and continue with CONF1 and CONF2. Section 6.2.2 presents the resulting design of experiment 2.

6.2.1 Changes of design and setup

CONF3**.** As mentioned earlier in section 5.4, my informants suggested that if the numeral in the target sentence was instead modified by *approximately* (*ongeveer* in Dutch), the target sentence would sound natural and would be more preferred than the actual target sentence with *minstens* ('at least') given the particular speaker ignorance contexts. Hence, assuming that the difference in question transfers to the relevant English counterparts, we need to ask what is it that makes *at least* more specific than *approximately* given an ignorance context that entails underspecified speaker ignorance?

A1 The core meaning of *at least*.

A2 The derivation of the pragmatic meaning associated with *at least*.

Let us first consider A1 and see in what sense the core meaning of *at least* is specific. By using *at least n* the speaker asserts that the quantity in question is in a certain range, i.e., in $[n, ...)$, hence she is competent and specific about

which particular values are excluded (i.e., those that are lower than n), which is not applicable or clear in the case of uttering *approximately n*, whereby the speaker just gives an estimation and the listener cannot tell precisely where the speaker's knowledge starts and ends. That is to say, in the former case a speaker uttering the target sentence of our experiment could be understood as being specific and knowledgeable to some extent because she excludes certain values.

What happens if we further take into account the information conveyed by the context preceding the target sentence in experiment 1? Assuming the specificity due to *at least*'s core meaning (A1), while being completely compatible with the authority contexts as well as with the ignorance contexts, the target sentence with *at least* might be interpreted as being too specific in the latter case (hypothesis H1). That is, it might be perceived as giving too much information compared to what was expected given the speaker's prior underspecified ignorance utterance, where she literally states that she is not aware of the exact quantity under discussion, but she does have an impression or opinion, or—in some other items—that she can give an estimation or say what she thinks. Consequently, the Context effect attested in the probability of re-reading the region of the numeral phrase (number+noun, Region 6) in experiment 1 is an indication of the following: Readers are likely to revisit that region because at some point during the incremental interpretation of the target sentence and after reading that region they realize the infelicity of the supposedly more specific and detailed information expressed by the *minstens* modified numeral.

Let us now consider A2, which states that another reason why an *at least*-utterance might be interpreted as giving specific information and the relevant speaker as being rather competent is the derivation of the pragmatic meaning associated with *at least*. As we have seen in section 2.2, the ignorance reading most of the pragmatic accounts derive for *at least n* is what I have called specific speaker ignorance: *The speaker considers it possible that n and she considers it possible that* $> n$. This reading is more specific than the underspecified ignorance reading where the speaker lacks any beliefs about the values in the range $[n, ...)$ (hypothesis H2), which is what is contextually entailed in the ignorance Context condition of experiment 1. Therefore, this is a second scenario where the information expressed by the target sentence in experiment 1 is compatible with, but more specific than, that expressed by the preceding ignorance context.

This is how we could then explain the attested processing penalty in experiment 1 if we assume H2: Participants derive (rapidly) the specific speaker ignorance inference when first reading the critical regions and then consider it infelicitous given the speaker's limited competence signaled by the ignorance context sentence. Note that in such a scenario participants most possibly derive the ignorance implicature not because they are prompted by the relevant ignorance context, but because they are not discouraged to do so, as is the case in the authority contexts. To be more specific about the Context effect of experiment 1 observed not in early measures but in the probability of re-reading,

a scenario like H2 has it that at first sight (i.e., early reading measures) the (rapidly) computed ignorance implicature possibly looked compatible with the preceding ignorance context, especially given that there was no explicitly contradicting cue provided by the context—as in the case of the authority context sentences, and on second or further thought (cf. re-reading probability), participants found the derived pragmatic information too specific given the context they had read.

To sum up, setting aside at this point CONF1 and CONF2 as additional possible confounding factors, the observed processing penalty in experiment 1 might have been due to:

(H1) the infelicity of the specificity of the core meaning of *at least* given the prior underspecified ignorance utterance of the speaker in the ignorance Context condition

(H2) the infelicity of the specificity of the specific ignorance implicature *at least* is associated with given the prior underspecified ignorance utterance of the speaker in the ignorance Context condition

(H3) the on-line computation of the specific speaker ignorance inference triggered in the target sentence with support by the ignorance context

In H3 we see the explanation I pursued in chapter 5, which was afterwards questioned by positing a number of alternative interpretations of the effect of interest.

Crucially, in H2 and H3, the presence of an ignorance implicature is presupposed. In order to determine whether the effect of interest is to be associated with the (rapid or on-line) derivation of an ignorance implicature or not, we need to add a condition that can tell the two possibilities apart (H2/H3 vs. H1). I propose to do so by also testing the comparative counterpart of the superlative numeral modifier under investigation.

More than m has the same core meaning as *at least m+1*, both setting the lower bound of a certain range of values and excluding the same values, e.g., the number n of souvlakis Magda ate in either case in (1) is in the range $[4, ...)$, assuming only discrete numbers of souvlakis.

(1) Magda ate at least 4/more than 3 souvlakis.

Given that, the comparative and superlative modifier do not differ with respect to A1 specificity.

Nevertheless, they do differ as far as A2 specificity is concerned if we take the following empirical evidence into account and make certain assumptions. Westera & Brasoveanu (2014) found in their first experiment that comparative modified numerals trigger less robust speaker ignorance inferences compared to superlative modified numerals when there is a *how many* question under discussion and superlative modified numerals are associated with a later pro-

cessing cost (five words after the numeral modifier).[1] Moreover, in chapter 3, specifically in section 3.4.2, where we were concerned with the experimental investigation of modal variation inferences of modified numerals, we gained insight into the type of implicature mechanism and of alternatives modified numerals are associated with. We showed experimentally that the relevant inferences include the inference of the possibility of the lowest number compatible with the modified numeral, that is, the implication given for (2) below.

(2) Magda has to eat at least 4/more than 3 souvlakis.
$\rightsquigarrow$ *Magda may eat exactly 4 souvlakis.*

We further demonstrated that this inference is more robustly associated with *at least* than with *more than* (actually, we tested their Dutch equivalents *minstens* and *meer dan*, respectively, and once more I assume that the attested difference transfers to English). Assuming that the same pragmatic mechanism is responsible for the derivation of variation and ignorance inferences (cf., e.g., Nouwen, 2015), a similar contrast is to be found with respect to ignorance inferences, namely, that the implication *the speaker considers it possible that Magda ate exactly 4 souvlakis* for (1) is more robust with *at least* than with *more than*. Actually, most possibly, this corresponds and is to be linked to the aforementioned contrast Westera & Brasoveanu (2014) find between the two types of numeral modifiers in implicature rate and in reading times in their first experiment in the *how many* QUD condition. Remember also that the inference about the minimum value compatible with the modified numeral has been experimentally shown by Mendia (2016c) to be a necessary epistemic possibility for an ignorant speaker making an (unembedded) *at least n*-utterance, or in other words, to be a necessary componenet of the resulting speaker ignorance inference. Note that Mendia (2016c) did not test *more than* though.

To conclude, we maintain that the comparative and superlative modifier differ (i) in the strength of the specific ignorance implicature they trigger, and (ii) with respect to the specificity due to the derivation of the specific ignorance implicature (A2) because of (i). Thus, given this difference between the comparative and the superlative modifier and their similarity with respect to A1 specificity, the comparative modifier *meer dan* ('more than') will be added to the design of experiment 2 enabling us to test whether the effect observed in experiment 1 is to be associated with the derivation of a specific speaker ignorance implicature. By extension, this will help us shed light on the incremental interpretation of numeral modifiers and the implicature-mechanism they involve. As a sneak preview of the relevant predictions, if H1 holds, we expect to see no difference between superlative and comparative items, since they exhibit no difference with respect to A1, while, if H2 or H3 hold, we do expect to see a difference, since these hypotheses involve the pragmatics of modified numerals, which is where the superlative and comparative modifiers do differ.

[1] No difference in ignorance rate was found with other precise QUDs or other types of QUD in this experiment.

In section 6.3, I am describing these predictions in great detail referring to the modification of the experimental design and to the discussed similarities and differences between the two types of modifiers.

I now turn to the other two confounds, i.e., CONF1 and CONF2. More specifically, in the following, I will be presenting the modifications we made to the texts tested in experiment 1 in order to deal with the confounding factors CONF1 and CONF2.

CONF1**.** To deal with CONF1, only round and imprecise numbers were used in the target sentences, following Jansen & Pollmann's (2001) definition of roundness in Dutch combined with Dutch speakers' intuitions as far as preciseness is concerned.

Jansen & Pollmann tested in a corpus study in Dutch which numbers appear in approximation contexts following the word *ongeveer* ('approximately'). Based on their findings, they concluded that a number is round if divided by powers of 2, 5, or 10 it outputs 1, 2, 3, 4, 5, 6, 7, 8, or 9. So we modified the numbers in the modified numeral phrase in the target sentence so as to be compatible with this conclusion. (3) is an example of a relevant modification we made.

(3) **Target sentence**

Experiment 1

Jim heeft minstens vierentwintig pakketten in de avond afgeleverd.
'Jim delivered at least **twenty-four** parcels in the evening.'

Experiment 2

Jim heeft 's avonds minstens vijfentwintig pakketten aan klanten afgeleverd.
'In the evening, Jim delivered at least **twenty-five** parcels to clients.'

Regarding preciseness, we modified the number modified in the target sentence so as to be a round number that was also imprecise when the granularity level set by the noun it combines with was taken into account (i.e., coarse granularity). Below you see such a modification we made.

(4) **Target sentence**

Experiment 1

Linda is minstens twee en een halve kilo in een maand afgevallen.
'Linda lost at least **one and a half** kilo in a month.'

Experiment 2

Linda is de afgelopen maand minstens twee kilo's met veel afzien afgevallen.
'Last month, Linda lost at least **two** kilos with a lot of effort.'

The exclusive use of round and imprecise numbers in the target sentences aimed to make them (more) natural and coherent utterances of the ignorant

speaker in the ignorance Context condition eliminating the relevant confounds.

CONF2. A new sentence was added in the Intro of the texts, which introduces an unmodified numeral phrase that relates to the modified numeral phrase to be mentioned in the target sentence and to the quantity that is being under discussion in each text; see, for instance, in (5) an example of how the Intro from experiment 1 was modified in experiment 2.

(5) **Intro**

Experiment 1

Wesley runs his own parlor, which he enjoys a lot. He always makes real works of art and it is usually very busy. That's why he wanted to work hard all through Thursday.

Experiment 2

Wesley runs his own tattoo parlor, which he enjoys a lot. It's usually very busy and he tries to tattoo eight customers per day. Last Thursday, he was very busy.

The idea of such a modification was that introducing a number early in the text would make the later mention of a number (i.e., the numeral in the target sentence) concerning the same matter more natural and relevant given a context with an ignorant speaker by creating some comparison or contrast between the two numeral phrases. By that means I aimed to get rid of confound CONF2.

6.2.2 New design

Having motivated the changes that were thought to be in order after obtaining the results of experiment 1, we shall now see what the precise design of experiment 2 looks like. We have again two manipulations, but now the manipulation of the Verb type of experiment 1—which did not yield interesting results—has been replaced by that of the type of numeral modifier. That is, we manipulated:

(i) the numeral modifier in the target sentence (NM manipulation): *minstens* 'at least' vs. *meer dan* 'more than'.

(ii) the speaker's epistemic state as set up by the context before the target sentence with the numeral modifier (Context manipulation), as in experiment 1: *speaker ignorance* vs. *speaker authority* context. We also included a third condition with a *speaker indifference* context, which we will not be concerned with in the present chapter, but rather in chapter 7.

Thus, the NM and Context factors were manipulated in a 2×3 design. Below I am giving an example of a test item, the one that corresponds to the example item from experiment 1, in all relevant four conditions. As in experiment 1, the texts consisted of an Intro followed by the context sentence with the Context

manipulation, which was in turn followed by the target sentence with the NM manipulation; some of the texts would end with an Outro (see details in section 6.5.2).

(6) **Example item**

Intro
Wesley heeft zijn eigen zaak waar hij met veel plezier tatoeages zet. Het is er meestal erg druk en hij probeert elke dag acht mensen te tatoeëren. Deze donderdag was hij ook weer hard aan het werk.
'Wesley runs his own tattoo parlor, which he enjoys a lot. It's usually very busy and he tries to tattoo eight customers per day. Last Thursday, he was very busy.'

Context sentence

Speaker ignorance context
Ik weet niet precies hoe het met de drukte zat, maar ik heb wel een idee.
'I don't know exactly how busy it got, but I have got an impression.'

Speaker authority context
Ik weet precies hoe het met de drukte zat en daarom zal ik je erover vertellen.
'I know exactly how busy it got, and I'll tell you about it.'

Target sentence
Wesley heeft die dag $\left\{\begin{matrix}\text{minstens}\\ \text{meer dan}\end{matrix}\right\}$ *tien mensen met veel oog voor detail*
Wesley has that day $\left\{\begin{matrix}\text{at least}\\ \text{more than}\end{matrix}\right\}$ *ten people with much eye for detail*
getatoeëerd.
tattooed.
'That day, Wesley tattooed at least/more than ten people with a real eye for detail.'

Outro
Hij heeft ook twee medewerkers in dienst die de kunst van het tatoeëeren van hem hebben geleerd.
'He also has two employees in his service that have learned the art of tattooing from him.'

Apart from the modification of the Intro and that of the number in the target sentence, some new material has been inserted between the verb *heeft* (which is no longer manipulated) and the numeral modifier, that is, the new phrase *die dag* 'that day' in (6). This addition was considered necessary because of information structure after introducing the contrast between the two numeral phrases (that in the Intro and that in the target sentence). To that end, in most of the items the temporal adverbial or prepositional phrase that served as a spillover region in the items of experiment 1 was fronted, because it contained

new information, and sometimes it was also modified according to the discoursal needs (cf. experiment 1: *on Thursday* → experiment 2: *that day*). New material had to be placed then in the spillover region, i.e., after the modified numeral phrase, cf., e.g., *met veel oog voor detail* ('with a real eye for detail') in the target sentence in (6).

6.3 Predictions

In section 6.2, I presented in detail the new experimental design and the various hypotheses this sets out to test, see also table 6.2. The present section discusses the predictions of each of the hypotheses in question.

Hypotheses re Context effect in experiment 1	**Predictions**
H0 not due to CONF1 or CONF2	Context effect in *ten people* in re-reading probability
H1 due to infelicity of specificity of core meaning of *at least* given prior underspecified ignorance utterance	No interaction effect in *ten people* in re-reading probability
H2 due to infelicity of specificity of specific ignorance implicature given underspecified ignorance utterance	Interaction effect in *ten people* in re-reading probability OR Different processing profile for the two NMs
H3 due to taxing online computation of specific ignorance implicature	Interaction effect in *ten people* in re-reading probability OR Different processing profile for the two NMs

Table 6.2: Hypotheses and predictions of eye-tracking experiment 2.

First, if we assume that the positive effect of Context we found in eye-tracking experiment 1 (which only tested the superlative modifier) was not due to CONF1 or CONF2, then we expect to replicate that effect in the present experiment (see H0 on table 6.2), as the Context manipulation remained the same as in experiment 1 (i.e., the contrast between authority and ignorance

Context conditions).

Let us see now what is predicted under H1, namely, if we assume that the Context effect observed in experiment 1 in the probability of re-reading measure is caused because of the core meaning of *at least* being perceived as too specific given the prior underspecified ignorance utterance in the ignorance Context condition. Under H1, we predict no interaction to take place between the NM and Context factors, but only a positive Context effect as in experiment 1. That is, a Context effect should arise irrespective of the type of the numeral modifier, as the core meaning of the superlative lower-bound modifier and the comparative lower-bound modifier is equally specific. Moreover, one could specifically expect the disruptive Context effect due to the resulting infelicity to appear in early measures rather than in late measures (which was the case in experiment 1, cf. probability of re-reading measure), since semantic interpretation is argued to happen rapidly and to be evident early in the eye-tracking record (see, e.g., Clifton et al., 2007). However, it is also likely that readers rapidly derive the core meaning of the modified numeral phrase and only later during incremental interpretation do they realize that this is infelicitous to some extent given what they have read in the ignorance context sentence. In such a scenario this realization can be manifested by a higher likelihood of revisiting the critical region with the numeral phrase, where the interpretation of the whole modified numeral phrase is completed (cf. effect in experiment 1).

If, on the other hand, we assume H2, where the specificity of the superlative modifier has to do with the derivation of the specific ignorance implicature, the predictions are very different. Given that the comparative modifier triggers the specific ignorance implicature to a smaller extent compared to the superlative modifier (see extensive discussion in section 6.2.1), we expect the following: The resulting infelicity of the specificity of the target sentence in the superlative NM condition given the prior underspecified ignorance utterance will be stronger than the corresponding infelicity in the comparative condition. As a result, the disruptive effect of that infelicity is expected to be manifested (i) as a stronger Context effect for *minstens* than for *meer dan* in the same region and measure (interaction of NM and Context factors), or (ii) by a different processing profile for the two NMs, if the strength difference of the respective ignorance implicatures is the result of a difference in the pragmatic mechanism involved and pragmatic mechanisms of a different nature are put into operation at different points of the eye-movement record, depending on how tied they are to the lexical meaning of the implicature-trigger. Given that, the manifestation of the infelicity under consideration might differ for the two NMs accordingly (i.e., depending on when their ignorance implicature was calculated).

Finally, the same predictions hold, if we consider H3. That is, if we assume that experiment 1's Context effect was not triggered due to the infelicity of the (A1/A2) specificity of the target sentence given the ignorance context sentence, but rather because of the taxing on-line calculation of the specific speaker ignorance implicature (as has been found for the on-line derivation of other types of (Quantity) implicatures by studies using a similar context setup

in a reading experiment, e.g., Breheny et al., 2006; Panizza et al., 2009, see section 5.3.3.1), an interaction between Context and NM is predicted in the same region and measure. That is to say, due to its lower robustness (because of the lower robustness of the exhaustive interpretation of the minimum value n of the relevant range), the specific ignorance implicature associated with the comparative numeral modifier is predicted to require less cognitive effort and, thus, manifest itself in on-line processing as a weaker Context effect for *meer dan* compared to *minstens*. Alternatively, we could observe a different processing profile for the two NMs, assuming that the strength difference of their ignorance implicatures is due to a difference in the pragmatic mechanism employed, which has a different on-line manifestation depending on how tied to the lexical meaning of the trigger it is.

As has already been mentioned, H2 and H3 can be subsumed under the hypothesis that the Context effect observed in experiment 1 is to be associated with the derivation of a specific ignorance implicature. I will use HH2 to refer to that hyper-hypothesis and HH1 to refer to the hypothesis that the Context effect is not to be associated with the derivation of a specific ignorance implicature, as in the case of H1 (or of CONF1 and CONF2). Table 6.3 lists those hyper-hypotheses repeating the relevant predictions from table 6.2.

	H-Hypotheses re Context effect in experiment 1	**Predictions**
HH1	Not associated with derivation of specific ignorance implicature	No interaction effect in *ten people* in re-reading probability AND Similar processing profile for the two NMs
HH2	Associated with derivation of specific ignorance implicature	Interaction effect in *ten people* in re-reading probability OR Different processing profile for the two NMs

Table 6.3: Hyper-hypotheses and predictions of eye-tracking experiment 2.

6.4 Off-line tests

In order to test how acceptable and coherent our constructed items are, especially after making the small changes in their Intro and the (information) structure of the target sentence, we collected off-line judgements from native

speakers of Dutch by means of a questionnaire (test 2) that resembled pretest 2 of experiment 1.[2] We also tested how strong our manipulated contexts are in terms of speaker's knowledgeability of the quantity under consideration; I am discussing this latter test first in the following section. Both tests were created in Ibex and hosted on Ibex farm (Drummond, 2007).

6.4.1 Test 1

In the first questionnaire, we tested the strength of the contexts in terms of knowledgeability in a different way from that used in pretest 1 of experiment 1. We specifically asked native speakers of Dutch (N=32) to judge the author's (of the texts) knowledgeability of the quantity being discussed. Participants were given 39 test items (Latin square design) intermixed with 36 filler items (75 trials in total), with the difference that the modified numeral in the target sentence was replaced with the hash symbol (#), as in (7) for the example item in (6).

(7) **Example item: Target sentence**
Wesley heeft die dag # mensen met veel oog voor detail getatoeëerd.
Wesley has that day # people with much eye for detail tattooed.
'That day, Wesley tattooed # people with a real eye for detail.'

They were told that # stands for a quantity and were next asked to indicate how likely it is that the author knows the precise quantity. They had to do so on a Likert scale from 1 to 7, where 1 stands for "highly unlikely" and 7 for "highly likely". The author of ignorance contexts was expected to receive low likelihood ratings, below the middle value of the scale, while the author of the authority contexts was expected to be judged as being highly likely to know the precise quantity. Apart from these two (and the indifference Context) conditions, we also included a condition where there was no context sentence, but the target sentence followed immediately after the Intro part. In that condition, we expected to see a split/chance performance, as the relevant context before the target sentence was not (explicitly) specified in terms of speaker's knowledgeability of the quantity being discussed.

The author in the ignorance contexts was judged to be highly significantly less likely to be knowledgeable of the precise quantity compared to the author in the authority contexts ($t = -9.022$, $p < .0001$). The latter condition scored very close to the right end of the relevant scale ($mean = 6.050$), while the former condition obtained a mean likelihood rating lower than the middle value of the Likert scale ($mean = 3.411$), as expected. Perhaps one would expect that the ignorance context condition would get an even lower mean score. However, the fact that the second half of the ignorance context sentences would state that the author does have an impression or suspicion or some thoughts with

[2]The results I am reporting on in the present chapter have also been included in a paper that appeared in SALT 27 proceedings as Alexandropoulou, Dotlačil & Nouwen (2017).

respect to the number in question possibly justifies the actual mean score of the condition. As to the condition where there was no context sentence intervening between the Intro and the target sentence, this received rather high scores, i.e., *mean* = 5.592, revealing that the Intro of the test items sets up the author/speaker as an authority on what she is talking about. Recall that an ignorance Context effect was attested in the Intro in eye-tracking experiment 1 in re-reading probability: Readers were found to be more likely to go back to the Intro after having read subsequent parts of the text, including the context sentence, when that sentence had an ignorant rather than an authoritative speaker. There, we hypothesized that this might be because the Intro creates a knowledgeability bias that is revisited after participants encounter an ignorance context sentence.

We conclude that the speaker authority contexts of experiment 2 are strong knowledgeability contexts with respect to the quantity being discussed, while the speaker ignorance contexts are poor(er) knowledgeability contexts.

6.4.2 Test 2

The second questionnaire tests how coherent our test items are and the actual task was as in pretest 2 of experiment 1. Participants (N=18) were given texts like (6) (without Outro) in all six conditions (Latin square design), repeated below as (8) in the same four conditions of interest, and were asked to judge how compatible the target sentence with the numeral modifier is with the (whole) preceding context.

(8) **Example item**

Intro
Wesley heeft zijn eigen zaak waar hij met veel plezier tatoeages zet. Het is er meestal erg druk en hij probeert elke dag acht mensen te tatoeëren. Deze donderdag was hij ook weer hard aan het werk.
'Wesley runs his own tattoo parlor, which he enjoys a lot. It's usually very busy and he tries to tattoo eight customers per day. Last Thursday, he was very busy.'

Context sentence

Speaker ignorance context
Ik weet niet precies hoe het met de drukte zat, maar ik heb wel een idee.
'I don't know exactly how busy it got, but I have got an impression.'

Speaker authority context
Ik weet precies hoe het met de drukte zat en daarom zal ik je erover vertellen.
'I know exactly how busy it got, and I'll tell you about it.'

Target sentence

Wesley heeft die dag $\left\{\begin{matrix}\text{minstens}\\ \text{meer dan}\end{matrix}\right\}$ *tien mensen met veel oog voor detail*

Wesley has that day $\left\{\begin{matrix}\text{at least}\\ \text{more than}\end{matrix}\right\}$ *ten people with much eye for detail*

getatoeëerd.

tattooed.

'That day, Wesley tattooed at least/more than ten people with a real eye for detail.'

Participants had to give their judgements on a Likert scale from 1 to 7, where 1 stands for "the sentence does not fit with the preceding context" and 7 stands for "the sentence fits well the preceding context". The idea was to exclude from the eye-tracking testing those items that would receive low mean coherence ratings overall. We tested the same thirty-nine items as before interspersed with forty-two filler items.

Given our participants' judgements, we excluded three items that received the lowest mean coherence ratings overall (< 4.25). The remaining 36 items (a number suitable for our 3×2 Latin square design), which were judged to consist in coherent texts (mean>4, the middle value of the Likert scale), were included for testing in the eye-tracking experiment.

As in pretest 2 of experiment 1, we statistically analyzed the overall dataset and we found a significant difference in coherence between items with speaker authority contexts and items with speaker ignorance contexts. More precisely, speaker ignorance context items received significantly lower ratings ($mean$ = 4.517) than authority context items ($mean$ = 5.509, $p < .01$). Although I will not discuss in detail the results regarding the speaker indifference Context condition until chapter 7, I can note here that this condition performed in a very similar way to the ignorance Context condition. No other effect or interaction was attested.

As in experiment 1, where the relevant highly significantly different means were 4.243 and 5.202 for the ignorance and the authority Context conditions, respectively, we did not include any bad/incoherent controls expected to score at the bottom of the coherence scale, and all 42 filler items, which were test items of other experiments, were highly coherent and well-formed texts (means>4, with the exception of two items that had mean overall score 3.389 and 3.722).[3] In the absence of ill-formed, clearly incoherent control items and due to the nature of the task that draws participants' attention to the coherence relation between the target sentence and the preceding text, it is very likely that participants became fastidious and stricter in their judgements, especially considering that they are given a scale of seven points to make some use of. As I remarked

[3]The filler items consisted of five types of items with the following mean scores across lists: 5.424 (min: 4.375, max: 6.625), 5.722 (min: 4.625, max: 6.875), 4.993 (min: 3.625, max: 6.750), 5.625 (min: 4.500, max: 6.875), and 5.439 (min: 3.2, max: 5.9).

in experiment 1, this might have resulted in the enlargement of possible small differences between items, hence the differences we observed across conditions.

Another possibility, however, could be that the difference we observe is due to the infelicity of the specificity of the target sentence given the prior less specific/underspecified ignorance utterance of the speaker according to some informants from our previous study, that is, it might be due to the confounding factor CONF3 discussed earlier in this chapter. Note that we did not make any changes (in that respect) on the ignorance or authority context sentences.

One could possibly even pursue a (rather far-fetched) scenario that was suggested to me whereby the finding in question resembles findings from the garden-path literature. In that literature, it has been shown that the taxing processing operations that take place in garden-path sentences make participants perform poorly on acceptability judgement or question-answering tasks (e.g., MacDonald, Just & Carpenter, 1992). By analogy, in such a scenario, the attested ignorance Context effect could be taken to reflect that participants are faced with an effortful cognitive process in the relevant condition such as the calculation of the specific speaker ignorance implicature, which is dictated by the corresponding contexts. That is, this effortful process taking place in the ignorance items causes participants to fail to perform the felicity judgement task in the ignorance condition as well as they do in the computationally troubleless authority condition. However, in such a scenario, we would additionally expect a difference between the two NM conditions (i.e., interaction effect), as comparative modifiers are taken to trigger weaker ignorance implicatures with a *how many* QUD (cf. Westera & Brasoveanu's (2014) results of first validity judgement task; Ciardelli et al., 2017) or no ignorance at all according to Coppock & Brochhagen (2013a); Geurts & Nouwen (2007); Nouwen (2010); Kennedy (2015) (see relevant details in section 2.2).

Luckily, this time we managed to shed some light on what kind of differences were occurring between different Context conditions. After the completion of the questionnaire we asked for participants' comments regarding (some of) their low item ratings. Interestingly, a reoccurring comment we received was that when the context had an ignorant speaker the target sentence—irrespective of the numeral modifier—felt like giving specific and detailed information that was in conflict with what the relevant speaker had just proclaimed. This is apparently in favor of our CONF3-story (i.e., H1 and H2) and, since no interaction was observed revealing that the resulting infelicity due to the specificity of the target sentence is weaker with the comparative modifier, it speaks in favor of hypotheses H1 and HH1. Namely, that the observed effect is due to the infelicity of the equally specific core meaning of the comparative and the superlative modifiers given the preceding ignorance utterance by the same speaker and it is not to be associated with the derivation of a specific ignorance implicature.

Adopting such an interpretation for the Context effect in the present questionnaire would mean that the infelicity predicted under H1/HH1 would have to be stronger than the clash and infelicity predicted to arise with *minstens* in the authority contexts according to the accounts that derive an obligatory speaker

ignorance inference for superlative modifiers, i.e., Ciardelli et al. (2017); Cohen & Krifka (2014); Coppock & Brochhagen (2013a); Geurts & Nouwen (2007); Nouwen (2010); Spector (2015); Spychalska (2015), see sections 2.2 and 5.3.2.1. Let us have a look at the mean scores per condition in figure 6.1 and evaluate these accounts.

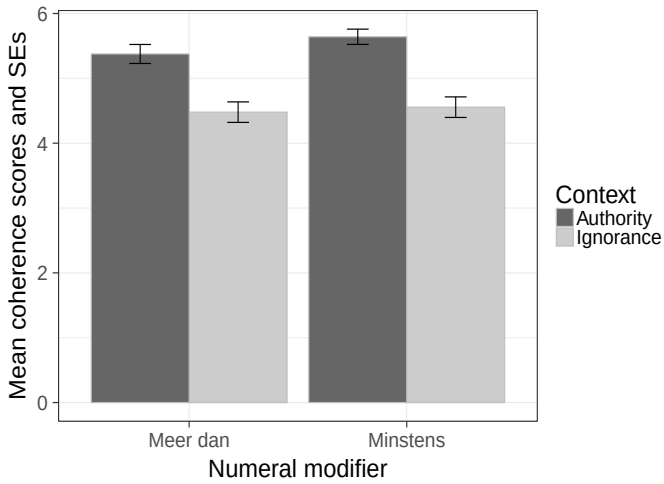

Figure 6.1: Mean compatibility scores of target sentence with preceding context per condition.

The results in figure 6.1 go against the accounts mentioned in the previous paragraph that derive an obligatory type of speaker ignorance. According to these accounts, the items with *minstens* in the authority Context condition are expected to receive low scores, because of the derivation of a strong, context-independent pragmatic or semantic implication of ignorance in the target sentence with *minstens* conflicting with the preceding authority context sentence. However, this condition exhibited the highest mean coherence score, i.e., 5.641, close to the right end of the Likert scale. Therefore, regardless of what these accounts predict about comparative modifiers (if they predict anything at all), what we see in figure 6.1 for *minstens* in authority contexts does not evince a severe pragmatic or a semantic inconsistency. The relevant finding, though, is in line with the neo-Gricean accounts by Büring (2008), Cummins & Katsos (2010), Kennedy (2015), Nouwen (2015), Schwarz (2016a), because they do not make any specification that would impose a context-independent and thus obligatory speaker ignorance implicature for superlative modifiers, which should then also arise in the speaker knowledgeability contexts of the present experiment. So once again, despite the factor that causes the ignorance Context condition to score lower than the authority condition, the present questionnaire reveals results that go against a large number of accounts of speaker ignorance

with *at least*.

Before moving to the actual experiment, I would like to point out that it is not necessary that the effect of test 2 will be reflected in the reading task of the eye-tracking experiment 2 or that it is to be connected with the Context effect in experiment 1 (see discussion in section 5.3.4.2 of chapter 5). Indicatively, we did not find any effect in participants' response times in the pretest in question that could be associated with the effect observed in the relevant judgement task (and, in fact, there was no effect whatsoever), though the length of the context sentences has not been strictly controlled for. Also, we should not forget that this off-line task was explicitly asking participants to focus on and judge the coherence relation of the target sentence with the preceding context. This is something that the reading task of the eye-tracking experiment does not do, except for the yes/no comprehension questions occasionally occurring after certain texts to control for participants' attention, which never ask anything related to the context or the target sentence.

6.5 Methods

6.5.1 Participants

Forty native speakers of Dutch participated in the eye-tracking experiment, who where recruited from the UiL OTS participant database. Two participants did not complete the experiment due to technical problems during testing, so the data of thirty-eight (4 male, mean age: 23.66, age range: 18–42) of the participants were further considered. All participants were paid (7 euro) for their participation in the experiment. Lastly, participants had normal or corrected to normal vision and were naive as to the purpose of the study.

6.5.2 Materials

We had thirty-six different texts, thirty-two of which were modified versions of the texts tested in experiment 1, while the remaining four items were newly constructed items. As already said, the texts consisted of an Intro, the context sentence, the target sentence, and occasionally of an Outro sentence (22 out of 36 items had an Outro). The Intro introduced the main character of the text and also a certain number to be picked up further. The Intro was the part that was modified the most, as texts had to be kept short after introducing the extra sentence containing the number. It consisted of two to five sentences and extended to three to five lines.

The context sentence that contained the Context manipulation extended to one up to two lines. The speaker ignorance and authority contexts were set up as in experiment 1 (see section 5.3.5.2; for the speaker indifference context sentences see sections 7.3.1 and 7.3.4 of the next chapter) and thus the content of the context sentences was left intact (except for tiny changes in the ignorance

condition of three texts and in the authority condition of one of them).

The target sentence had the NM manipulation (*minstens* 'at least' vs. *meer dan* 'more than') and was as in experiment 1 with the exception of the introduction of an extra phrase between the verb *heeft* 'has'/*is* 'is' and the numeral modifier for information structure reasons, as explained earlier in section 6.2. Most of the time, this additional phrase, constant across conditions, is a temporal prepositional or adverbial phrase (up to four words long). It was usually the phrase that served as a spillover region (following the modified numeral phrase) in the items of experiment 1. The numeral modifier (*minstens/meer dan*) combines with a round and imprecise number (see section 6.2), which is the same in both NM conditions and is given by one or two words. We used both low (e.g., *two, eight, fifteen*) and higher numbers (*forty, two hundred, three thousand, two million*). The numbers can be a numeral or a measure phrase.[4] The nominal material of the modified numeral phrase was from one to two words long and was identical across conditions. The numeral modifier as well as the number phrase are equally critical as the interpretation of the whole modified numeral phrase is completed after reading all this material (see section 6.6 re splitting into regions of interest). The new material placed in the spillover region (right after the modified numeral) was composed of two to five words (usually a prepositional or an adverbial phrase), constant across conditions and long enough to serve as such a region. The last word of the target sentence was always a participle of the main verb, the same in all different conditions.

Last, the Outro part, when present, consisted of one sentence that extended to one up to three lines.

The 36 test items appeared in six conditions (given our 3 CONTEXTs × 2 NMs design). The stimuli were rotated through twelve lists in a Latin square design and every list had six items per condition. Every participant saw only one list with 112 trials in total: Four practice items, thirty-six test trials, and seventy-two filler trials (see Appendix C). The trials of every list were randomly ordered for each participant. Fifty-two comprehension statements about the story narrated in the texts were included too, to control for whether participants pay attention to the texts they are reading. Participants had to indicate whether the statement was true or false given the text they had just read. These comprehension statements targeted information that was irrelevant to the manipulations of the experiment. Twenty-one (out of the 36) experimental items were followed by a comprehension statement, while the remaining 31 comprehension statements were follow-ups of the filler items.

Test items extended to six to eight lines and filler items between two and seven lines. Every line included up to sixty-eight characters. After the addition of an extra phrase, the target sentence had to be split into two lines most of the time. The line break most of the time occurred four words after the numeral modifier.

[4] As in experiment 1, I am assuming that there is no difference in *minstens/meer dan* modifying a bare numeral or a measure phrase.

6.5.3 Procedure

Participants were seated in a comfortable chair in a sound-treated booth and the distance of their head to the screen was 55–70 cm. Their eye movements (of their right or left eye) were recorded with an EyeLink 1000 eye tracker in remote mode (using a target sticker placed on participants' forehead), sampling at 500Hz. The stimuli were presented on a 17-inch Acer AL1717 monitor and a three-button button box was provided to participants for answering the comprehension questions or moving on.

Participants first read the instructions, where they were informed that they would be presented with short stories, each consisting of a short paragraph. After the story, a comprehension question about the text just presented would occasionally appear, which they had to answer by using the button box (right button for YES and left button for NO). Participants had to press the third, middle, button in order to progress. After reading the instructions, the calibration procedure with nine fixation points would start. After a successful calibration, participants would move on to the practice block, where they read four practice items and answered comprehension questions that two of them were associated with. Then the experiment block began. Before the presentation of a stimulus, a fixation point appeared on the screen to mark the beginning of the text and help participants find the start of the first sentence and avoid a seeking behavior that can influence the reading data. Participants were instructed to always fixate on this point, which would disappear once successfully fixated, giving its place to a story. Lastly, participants were instructed to read the stories at a normal pace, as they would do in their everyday life. The whole experiment lasted approximately 30-35 minutes.

6.6 Results

One of the 38 participants was excluded prior to analysis because she answered only 63% of the comprehension statements correctly, while the remaining 37 participants scored 77% or higher. The data of all remaining 37 participants were included in the statistical analyses. Data points of an item seen by 11 participants were excluded because it contained a typo in the context sentence (.86% of the observations). Data clean up was performed by using the software *Fixation* (Cozijn, 2006) and as described in eye-tracking experiment 1 (section 5.3.6). Another 1.12% of the overall observations was excluded because of problematic trials, that is, trials with few or no fixations due to track losses or participants' sloppy reading. For all measures, fixations of 0 ms were excluded from the analysis. We log-transformed the remaining reading-time data to approximate normality.

We split the texts into regions for the purpose of the analyses. As in experiment 1, the Intro constituted Region 1, while the speaker context was Region 2. The target sentence was divided into individual regions as illustrated in Table

6.4 for the example item in (7). The Outro of the texts was excluded from the region partitioning, as not all test items contained an Outro sentence.

Region 3	**Region 4**	**Region 5**	**Region 6**	**Region 7**	**Region 8**	**Region 9**
Wesley	has	that day	{more than / at least}	ten people	with much eye for detail	tattooed.

Table 6.4: Eye-tracking experiment 2: Regions of target sentence.

For each of the aforementioned regions we analyzed the same seven eye-tracking measures as in experiment 1: (i) *first-pass time* (sum of fixations in a region before exited for the first time to any direction), (ii) *right-bounded time* (sum of fixations in a region before exited for the first time to the right), (iii) *regression path duration* (sum of fixations since first fixation in a region until it is exited to the right for the first time), (iv) *probability of regression* (the probability of doing a first-pass regression from a given region), (v) *total reading time* (sum of all fixation durations in a region), (vi) *re-reading time* (sum of fixations in a region excl. first-pass fixations), and (vii) *probability of re-reading* (the probability of entering a region more than once). Tables 6.5, 6.6, and 6.7 report the raw means of the eye-tracking measures per region, including only values for regions that were fixated after reading later, critical, parts of the individual texts containing the Context and the NM manipulations.

The statistical analysis proceeded as in experiment 1. We conducted linear mixed-effects regression analyses on the log-transformed reading-time data and mixed-effects logistic regression analyses on the categorical regression probability and re-reading probability measures using the `lme4` package (Bates et al., 2015) in `R`. All analyses included two predictors, speaker Context (authority / ignorance / indifference context) and numeral modifier (NM, superlative vs. comparative), which were treatment-coded, with authority and superlative as the reference levels, respectively. The analyses also included random intercepts and slopes for participants and items for all fixed effects. We applied backward model selection for random effects (Barr et al., 2013) and in the following I will be presenting the output of the model with the maximal random-effect structure that converged and had the best fit. Apart from determining whether the speaker knowledgeability Context affects the incremental interpretation of modified numerals, we also aimed to determine whether it does so differentially for each NM condition. Therefore, the interaction of the two factors was generally kept in the models, even if it was not significant but had the expected sign (see Gelman & Hill, 2007, p. 69), cf. the predictions in section 6.3. *P*-values are calculated based on Satterthwaite's approximations using the *ImerTest* package of R. Results will be discussed per region, while the results for the indifference Context condition will not be reported until chapter 7 that deals with the speaker indifference implications of modified numerals.

Condition		Region		
Context	NM	Region 1 Intro	Region 2 Context	Region 3 *Wesley*
First pass (ms)				
Authority	comparative			279
Authority	superlative			303
Ignorance	comparative			242
Ignorance	superlative			265
Right bounded (ms)				
Authority	comparative			301
Authority	superlative			307
Ignorance	comparative			261
Ignorance	superlative			282
Regression path duration (ms)				
Authority	comparative			344
Authority	superlative			334
Ignorance	comparative			291
Ignorance	superlative			328
Probability of regression				
Authority	comparative		.897	.947
Authority	superlative		.874	.934
Ignorance	comparative		.945	.942
Ignorance	superlative		.887	.928
Total reading time (ms)				
Authority	comparative	6903		300
Authority	superlative	7072		325
Ignorance	comparative	6926		313
Ignorance	superlative	6965		299
Re-reading time (ms)				
Authority	comparative	1269		267
Authority	superlative	1235		311
Ignorance	comparative	1302		253
Ignorance	superlative	1337		274
Probability of re-reading				
Authority	comparative	.365	.266	.386
Authority	superlative	.341	.285	.418
Ignorance	comparative	.370	.228	.458
Ignorance	superlative	.405	.306	.441

Table 6.5: Eye-tracking experiment 2: Means of eye-tracking measures for regions 1, 2, and 3 in different conditions.

Condition		Region		
Context	NM	Region 4 *has*	Region 5 *that day*	Region 6 NM
First pass (ms)				
Authority	comparative	228	307	235
Authority	superlative	228	327	227
Ignorance	comparative	244	305	231
Ignorance	superlative	226	300	224
Right bounded (ms)				
Authority	comparative	243	371	246
Authority	superlative	241	385	246
Ignorance	comparative	251	366	241
Ignorance	superlative	232	387	243
Regression path duration (ms)				
Authority	comparative	315	464	277
Authority	superlative	330	476	320
Ignorance	comparative	313	453	276
Ignorance	superlative	301	496	290
Probability of regression				
Authority	comparative	.752	.760	.902
Authority	superlative	.700	.746	.860
Ignorance	comparative	.718	.717	.918
Ignorance	superlative	.726	.711	.894
Total reading time (ms)				
Authority	comparative	282	464	308
Authority	superlative	281	502	329
Ignorance	comparative	299	468	328
Ignorance	superlative	253	510	320
Re-reading time (ms)				
Authority	comparative	273	402	290
Authority	superlative	259	432	304
Ignorance	comparative	278	436	307
Ignorance	superlative	237	514	307
Re-reading probability				
Authority	comparative	.468	.452	.429
Authority	superlative	.503	.488	.402
Ignorance	comparative	.465	.422	.467
Ignorance	superlative	.463	.458	.376

Table 6.6: Eye-tracking experiment 2: Means of eye-tracking measures for regions 4, 5, and 6 in different conditions.

Condition		Region		
Context	NM	Region 7 *ten people*	Region 8 PP	Region 9 *tattooed*
First pass (ms)				
Authority	comparative	386	367	262
Authority	superlative	399	404	279
Ignorance	comparative	388	402	258
Ignorance	superlative	385	418	242
Right bounded (ms)				
Authority	comparative	444	390	319
Authority	superlative	438	442	330
Ignorance	comparative	427	469	297
Ignorance	superlative	429	459	294
Regression path duration (ms)				
Authority	comparative	550	636	903
Authority	superlative	532	631	1116
Ignorance	comparative	518	710	1077
Ignorance	superlative	537	641	1120
Probability of regression				
Authority	comparative	.794	.818	.479
Authority	superlative	.827	.827	.475
Ignorance	comparative	.793	.768	.521
Ignorance	superlative	.849	.800	.475
Total reading time (ms)				
Authority	comparative	544	460	343
Authority	superlative	517	528	348
Ignorance	comparative	549	543	311
Ignorance	superlative	564	527	315
Re-reading time (ms)				
Authority	comparative	421	314	360
Authority	superlative	432	394	313
Ignorance	comparative	471	446	311
Ignorance	superlative	499	368	287
Re-reading probability				
Authority	comparative	.429	.424	.270
Authority	superlative	.319	.426	.254
Ignorance	comparative	.396	.421	.224
Ignorance	superlative	.410	.453	.328

Table 6.7: Eye-tracking experiment 2: Means of eye-tracking measures for regions 7, 8 and 9 in different conditions.

6.6.1 Region 1, Intro

There was no main effect of NM or Context or an interaction effect of the two factors on any of the measures of interest in this region, i.e., total reading time, re-reading time, and probability of re-reading.

6.6.2 Region 2, Speaker knowledgeability context

As also explained in experiment 1, looking at the reading-time measures of the present region would not be any informative, as the two types of Context under consideration here (viz., authority and ignorance contexts) differ with respect to the lexical material used in order to set up the respective speaker's knowledgeability state and sometimes they differ in sentence length as well. Possible reading-time differences caused by such differences are uninteresting and irrelevant to the purposes of the present study.

The first measure of interest for the present region is the probability of regression, that is, in our case how likely it is that readers will regress back to the Intro. Recall that in pretest 1 (section 6.4.1), we found that the additional no-context Context condition received high speaker knowledgeability likelihood ratings ($mean$ = 5.592) and we concluded that this might indicate that the Intro sets up an author/speaker that is an authority on what she is talking about. Under this assumption, we could perhaps expect to already find an effect of the ignorance Context on regression probability. Such an effect would show that people are likely to revisit the Intro with the hypothesized authority bias after reading an ignorance vs. an authority context, because of the resulting inconsistency in the former case. No such effect was revealed by the statistical analysis, which included only Context as a predictor (authority: $mean$ = .115, ignorance: $mean$ = .0839, $z = -1.541$, $p = 0.123$), as at this point the region with the NM manipulation has not been encountered yet. No effect was found for the probability of re-reading measure either, the second measure under investigation in the present region.

6.6.3 Region 3, *Wesley*

As in experiment 1, there was no effect on any measure in the first region of the target sentence.

6.6.4 Region 4, *has*

The statistical analysis revealed no significant effect in Region 4, consisting of an auxiliary verb (*has/is*), in any of the measures. Recall that this was the case for experiment 1 too (except for some marginal effects), although there Region 4 contained the Verb manipulation (*has* vs. *had to*).

6.6.5 Region 5, *that day*

No main effects or interaction of Context and NM type on any measure were observed at Region 5, a region that was not included in experiment 1 between the auxiliary verb and the numeral modifier regions.

6.6.6 Region 6, NM

There was no main effect of or interaction between Context and NM type on any measure in Region 6, which contains the NM manipulation. Note that in experiment 1 there was no effect of Context in the region of the numeral modifier *minstens* ('at least') either.

6.6.7 Region 7, *ten people*

Similarly to experiment 1, we find effects on the probability of re-reading the critical Region 7, that is the numeral phrase (*ten people*) following the numeral modifier. More precisely, a significant positive main of Context (β = .446, SE = .216, z = 2.067, p < .05) was attested as well as an interaction between the Context and NM factors in that region in re-reading probability (β = −.560, SE = .299, z = −2.006, p < .05): The probability of re-reading Region 7 (*ten people*) of superlative items was higher in the ignorance than in the authority Context condition and this difference was bigger in the superlative than in the comparative condition, as also illustrated by Figure 6.2. That is, readers were more likely to re-read the critical region of the numeral phrase in an ignorance context as opposed to an authority context in the superlative condition and this happened significantly more in the superlative condition than in the comparative condition.

Importantly, the former observed effect of Context constitutes a replication of the Context effect attested in experiment 1 (see section 5.3.6.6), where readers were found to be more likely to re-read the same numeral phrase following the superlative modifier when they were in an ignorance as opposed to an authority context (in the non-modal Verb baseline condition). A post-hoc analysis for the separate numeral modifiers showed that the Context factor had no effect on the comparative NM condition though (β = −.154, SE = .207, z = −.742, p = .987).

In addition, as also manifested by Figure 6.2, the statistical analysis revealed a significant positive effect of NM on the probability of re-reading in the same region (β = .556, SE = .223, z = 2.489, p < .05), showing that readers were more likely to re-read that region when it followed *meer dan* ('more') as opposed to *minstens* ('at least') in an authority context.

No effect on any other measure was found to be significant in Region 7.

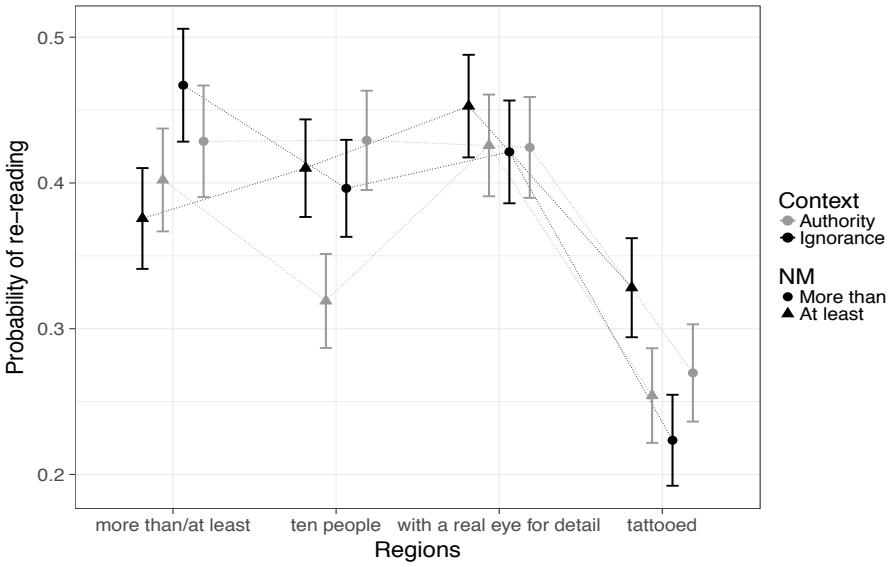

Figure 6.2: Probability of re-reading each region from the numeral modifier region onward per context condition.

6.6.8 Region 8, PP

In Region 8 (*with a real eye for detail*), there was a positive interaction of NM and Context factors (β = .456, SE = .180, t = 2.531, p < .05) on re-reading times. Post-hoc analyses for the separate numeral modifiers showed that there was an ignorance Context effect in the comparative NM condition (β = .272, SE = .107, t = 2.551, p < .05). These two effects taken together indicate that when readers re-read the prepositional phrase in Region 8 (see how often they re-read it in figure 6.2, also in relation to the superlative condition), (i) they exhibit a slowdown if they are in an ignorance context as opposed to an authority context in comparative items, and (ii) this slowdown is further bigger in comparative items than in superlative items (see figure 6.3). No significant effect on the other measures was found in Region 8.

6.6.9 Region 9, *tattooed*

Region 9 (*tattooed*), the last region of the target sentence, exhibited only a marginal negative interaction effect between NM and Context on re-reading probability (β = −.670, SE = .354, z = −1.890, p = .059), such that the probability of re-reading that region in an ignorance context as compared to an authority context is higher in the superlative than in the comparative condition. This difference is visible in Figure 6.2, too. Hence, once again, similarly

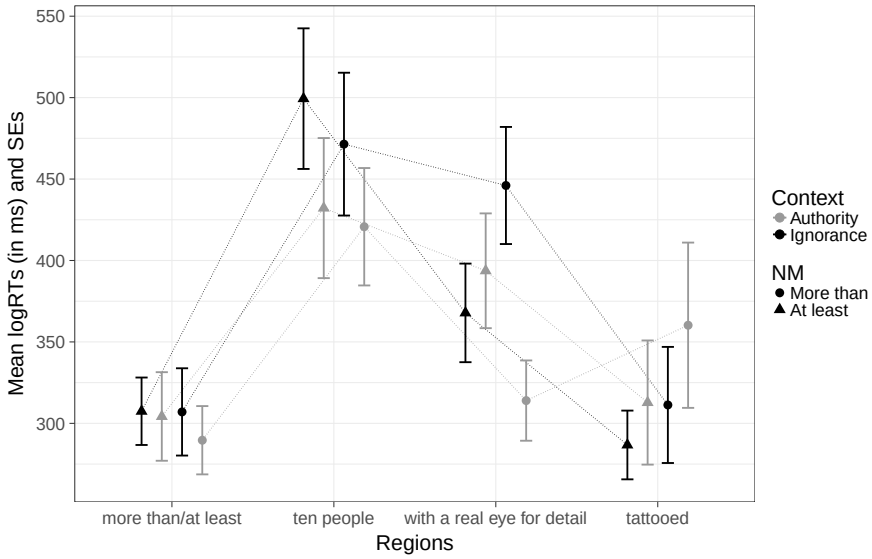

Figure 6.3: Re-reading times for each region from the numeral modifier region onward per context condition.

to experiment 1, no significant effect on any measure was found for the last region of the target sentence.

6.7 Discussion

Comparison of experiments 1 and 2. Overall, as far as the manipulation shared by experiments 1 and 2 is concerned, we see a similar picture in the two experiments: We find effects on the probability of re-reading the critical region with the numeral phrase following the numeral modifier at the expense of the superlative condition, while we do not find any new relevant significant effect throughout the various regions. The only exception is the effect of Context on re-reading probability attested in experiment 1 at the Intro (Region 1), which was not replicated in experiment 2. This might have to do with the fact that the Intro was considerably modified in experiment 2, after the need to introduce a number that relates to the quantity being discussed in the whole text. This further means that the effect in experiment 1 was not to be attributed to a possible knowledgeability bias of the Intro clashing with a following ignorance context sentence and making people revisit and reconsider the information in the Intro (see similar conclusion in pretest 1 of experiment 2), but rather it seems that it possibly had to do with the mention or choice of the number

being modified in the target sentence. That is to say, the introduction of a number in the Intro to be contrasted with that in the target sentence possibly eliminated participants' need to revisit the Intro in experiment 2.

But why? Why would readers revisit the Intro after encountering the modified numeral in the target sentence in the ignorance condition of experiment 1? It is likely that it was unexpected that the ignorant speaker would mention some number, even though she had just stated that she has got an impression or opinion or she is able to give an estimation. That is, the Context effect at the Intro of experiment 1 might be due to CONF2 (see table 6.8 below from section 5.4).

Confounds

CONF1	No control for preciseness/roundness of number modified by *minstens*
CONF2	Unexpectedness of use of a number in the target sentence
CONF3	Information of *minstens*-phrase too specific given underspecified ignorance from context

Table 6.8: Potential confounding factors of ignorance Context effect in Region 6 in eye-tracking experiment 1.

Another explanation of that effect could be that readers were wondering where this number came from and why this specific number was picked in the modified numeral phrase uttered by the particular ignorant speaker, making them seek further to find out whether and how this choice relates to anything the speaker had previously mentioned, for instance, in the Intro. Although these two interpretations of the Context effect in the Intro of experiment 1 disappearing in experiment 2 are very plausible, they fail to explain the interaction effect of the Context factor with the Verb factor of experiment 1 in re-reading time and probability in the same region. As there are indications of the Intro having a knowledgeability bias in experiment 2 too, due to the high mean knowledgeability score of the no-context condition in the off-line test 1 ($mean$ = 5.592, see section 6.4.1), it might be the case that the power of the present experiment is not enough for the relevant effect to show up. Note that the difference in question is in the right direction though (β = .319, SE = .225, z = 1.419, p = .156).

Different processing profile of superlative and comparative modifiers. We found that the processing profile of the two numeral modifiers differs: While

a processing penalty is detected at the numeral phrase following the superlative modifier when used by an ignorant speaker (in late measures, as in experiment 1), a processing penalty occurs at a later point of the target sentence with the comparative modifier uttered by an ignorant speaker (in late measures). This difference in the location of the attested effects points to a difference in the source and nature of those effects.

How does this finding compare to Westera & Brasoveanu's (2014) self-paced reading results for superlative and comparative numeral modifiers? Why do we find such a difference (specifically, positive effect of Context and negative interaction in Region 7 vs. positive interaction in Region 8), while (in their first on-line experiment) Westera & Brasoveanu found a similar processing pattern for their comparative and superlative items throughout the four words (*ten*, *of*, *the*, *diamonds*) following the modifier in the *how many* QUD condition and a difference at the expense of superlative items in the preposition (*under*) after the noun phrase of the modified numeral (five regions after the NM)? Of course, the context manipulation of the two experiments under comparison differs (while the structure of target sentences is similar). Nevertheless, we should perhaps link the absence of the relevant difference (i.e., absence of interaction) in Westera & Brasoveanu's (2014) experiment with the fact that in our experiment the relevant effects appear in late measures, e.g., re-reading probability for the superlative NM and re-reading time for the comparative NM. Importantly, those measures are not available in a self-paced reading task, where participants are not free to go back and re-read previously read parts upon an afterthought, for instance. This also demonstrates the advantage of an eye-tracking reading task over a self-paced reading task in providing (more) means for detecting subtle effects. On the other hand, one could maintain that we should identify the effect in re-reading probability for the superlative modifier (negative interaction effect) in our experiment with that appearing later in the target sentence of superlative items in Westera and Brasoveanu's experiment. Moreover, as this happens in the penultimate (or antepenultimate, depending on the item) region/word of the target sentence in Westera and Brasoveanu's experiment, it is possible that the following (and final, in some items) region (only one word) is not enough material for the manifestation of a processing penalty corresponding to that we found for our comparative items at a later point of the target sentence.

Region 7: Context and interaction effects. Going back to the motivation of the present experiment, our aim was to examine the possibility that CONF3 (see table 6.8) was a confound of the Context effect in experiment 1 and, by extension, find out whether that effect is to be associated with the derivation of the specific speaker ignorance implicature of the superlative numeral modifier (cf. hypothesis HH2 or sub-hypotheses H2 and H3 on table 6.9 combining tables 6.2 and 6.3 from section 6.3) or with the specificity of the core meaning of the superlative modifier being infelicitous in the context of the

preceding underspecified ignorance utterance (see H1 on table 6.9). So, given that its comparative counterpart has an equally specific core meaning and a supposedly less robust specific ignorance implicature, we included a comparative NM condition in experiment 2 in order to tease the above possibilities apart.

H-Hypotheses re Context effect in experiment 1	**Hypotheses re Context effect in experiment 1**	**Predictions**
	H0 not due to CONF1 or CONF2	Context effect in *ten people* in re-reading probability
HH1 Not associated with derivation of specific ignorance implicature	**H1** due to infelicity of specificity of core meaning of *at least* given prior underspecified ignorance utterance	No interaction effect in *ten people* in re-reading probability AND Similar processing profile for two NMs
HH2 To be associated with derivation of specific ignorance implicature	**H2** due to infelicity of specificity of ignorance implicature given underspecified ignorance utterance	Interaction effect in *ten people* in re-reading probability OR Different processing profile for two NMs
	H3 due to taxing online computation of specific ignorance implicature	Interaction effect in *ten people* in re-reading probability OR Different processing profile for two NMs

Table 6.9: Hypotheses and predictions of eye-tracking experiment 2.

We found the same positive Context effect as in experiment 1 in the region of the numeral phrase (*ten people*, Region 7): Comprehenders are more likely to reread the numeral phrase following the superlative modifier in an ignorance context than in an authority context. This effect confirms hypothesis H0, i.e., that that effect in experiment 1 has nothing to do with the roundness/preciseness (CONF1) or unexpectedness (CONF2) of the number being modified in the target sentence in the ignorance condition. We further found that that effect was stronger with the superlative modifier than with the comparative one (negative interaction effect). Given our hypotheses in table 6.9, the observed interaction

effect rejects H1 and HH1, and gives support to HH2 and sub-hypotheses H2 and H3. Thus, the Context effect replicated in experiment 2 is to be associated with the (rapid or on-line) derivation of the specific speaker ignorance implicature of the superlative modifier. That is, comprehenders revisit more often this specific region, where the interpretation of the whole modified numeral phrase is completed, (i) in order to derive the specific ignorance implicature, which is not entailed by the context, or (ii) because having already derived that implicature (rapidly) they realize afterwards that the resulting reading is too specific compared to the underspecified speaker ignorance entailed by the context. In other words, that effect can be taken to reflect the on-line derivation of the specific speaker ignorance implicature of *at least* or readers' real-time realization that the resulting ignorance reading from the implicature computation, *not* traced on-line, is too specific in the relevant context.

Moreover, as our context setup is the exact same as in experiment 1 and given that the off-line test 2 revealed high compatibility rates for target sentences with the preceding authority context sentences (*mean* = 5.509; specifically, *mean* = 5.641 for superlative condition) in experiment 2 too, the relevant results, taken together, go against the accounts that derive obligatory speaker ignorance with superlative modifiers such as Ciardelli et al. (2017), Cohen & Krifka (2014), Coppock & Brochhagen (2013b), Geurts & Nouwen (2007), Nouwen (2010), Spector (2015), and Spychalska (2015). On the other hand, our overall findings for the superlative modifier and their indication speak to the advantage of neo-Gricean Quantity-based accounts of (specific) ignorance implications such as Büring (2008); Kennedy (2015); Nouwen (2015); Schwarz (2016a).

Lastly, though the attested negative interaction effect and the absence of a Context effect in the comparative condition already reveal significant insight into the processing of comparative modified numerals (also in comparison to the superlative ones), even more can be said considering the effects observed in Region 8.

Effects in Region 8. We found that readers slow down when re-reading Region 8 (*with a real eye for detail*) of comparative items in the ignorance condition as compared to the authority condition and this slowdown is reliably larger in the comparative rather than in the superlative condition. As already said, what caused these effects is different from what caused the corresponding effects found in the previous region (in re-reading probability). We may interpret the effects in Region 7 to reflect the derivation of a specific ignorance implicature with *at least* in line with neo-Gricean accounts involving a Quantity-based mechanism. Or we may interpret the Context effect in Region 7 to be due to the infelicity of the specificity of the core meaning of *at least* given the prior underspecified ignorance utterance (if we think that the interaction effect is not telling). This means that the effects in Region 8 should not be given neither of these explanations. They should definitely not be connected

with the derivation of an ignorance implicature of the exact same nature or strength. Perhaps what we see in Region 8 is not even to be connected to any reasoning that derives ignorance implications. Perhaps what we find is an indication that a completely different type of pragmatic reasoning takes place. In other words, it is rather open what the effects in question reveal. But let us consider a number of possible scenarios.

First, the difference in the processing profile of the two NMs cannot be explained by the difference of the respective numeral modifiers in frequency. A corpus search I did showed that *meer dan* is considerably more frequent than *minstens* (194.678 vs. 31.381 hits, respectively), though the type of context was not taken into account (authority vs. ignorance).[5] Less frequent words have been found to be read longer compared to more frequent words and such frequency effects are argued to occur rapidly once the critical word is read and in early measures such as first fixation duration (i.e., duration of first fixation on a word) and gaze duration (i.e., the sum of all first-pass fixations on a word) (Clifton et al., 2007; Just & Carpenter, 1980; Rayner, 1977). Unless the frequency of the numeral modifiers varies depending on the type of context, the effects in Region 8 at the expense of the comparative condition should not be explained in terms of frequency. More importantly, the processing penalty associated with the superlative or the comparative condition occurred at a later region than the critical NM region, i.e., in Region 7 and 8, respectively, and also at a later point of the eye tracking record (late measures). Thus, neither of these effects could be a frequency effect indication.

Furthermore, we should not take the positive interaction effect in re-reading time in Region 8 to indicate a difference in the syntactic structure of comparative and superlative modified numeral phrases or in their integration in the syntactic representation. On such a syntactic interpretation, both the effect of ignorance Context and the interaction effect we find in Region 8 (in re-reading times) would mysteriously require a different structure for the target sentence with *unembedded* occurrences of numeral modifiers depending on the context. Also, both effects appear after the syntactic integration of the numeral phrase has been completed (Region 8: PP) and even appear in late measures, while, as revealed by Clifton et al.'s (2007) overview chapter, syntactic parsing is generally reflected at least in early measures of eye movements on the critical word/region or on the region after.

Effects in Region 8 due to ignorance implicature derivation (HH2). How can we then explain the attested processing penalty in the comparative items (in re-reading times) in Region 8, also in comparison with the superlative items? We could possibly take the slowdown in the ignorance Context condition

[5] I used the Corpus Hedendaags Nederlands (Corpus of Contemporary Dutch), which consists of 800.000 texts taken from newspapers, magazines, news broadcasts and legal writings (1814–2013). Specifically, this corpus is a combination of the 5, 27 and 38 Million Words Corpora and the PAROLE Corpus, supplemented with newspaper texts from NRC and De Standaard (until 2013).

of comparative items to be a signal of the derivation of (i) a weaker speaker ignorance implication, and/or (ii) a speaker ignorance implication with *more than* different in nature than that of *at least* that kicks in late(r) in the target sentence.

Specifically, as also discussed for *at least*, the observed processing penalty could be associated with the on-line computation of a specific speaker ignorance implicature (that includes the exhaustive interpretation of the minimum value of the expressed range), which is supported by the relevant context and QUD (H3). As pointed out earlier (see, e.g., section 6.2), such an implicature is weaker with comparative modifiers than with superlative modifiers. Hence, in those cases that participants re-read the PP region following the numeral phrase, they slow down in the ignorance condition of comparative items because only then do they compute the total interpretation of the modified numeral phrase including the specific ignorance inference. We take the location and timing of the relevant processing penalty to be an indication (I1) of the difference in implicature robustness between the two types of numeral modifiers and (I2) perhaps even of a difference in nature. Elaborating on (I1), the specific ignorance implicature is more loosely associated with the comparative modifier, hence it only occurs *if* readers re-read Region 8 in the ignorance condition, whereas in the superlative case readers are more likely to revisit the region of the numeral phrase (Region 7)—where the interpretation of the whole modified numeral is completed—in the ignorance condition in order to derive and establish the specific ignorance interpretation.

Ciardelli et al.'s (2017) account is an indicative case of a combination of I2 and I1. Ciardelli et al. derive a specific speaker ignorance implicature for both types of numeral modifiers, only as a Quality implicature with superlative modifiers and as a standard Quantity implicature with comparative modifiers (assuming a *how many* QUD for the latter). They further argue that in the former case the ignorance implicature is stronger because observing a Quality maxim is more urgent than observing a Quantity maxim.

Other than assuming H3, we can alternatively consider a specificity scenario, i.e., H2. According to such a scenario, the attested processing penalty in Region 8 is not to be associated with the on-line derivation of the specific ignorance implicature but rather with the on-line manifestation of the infelicity that the effortless derivation of that implicature at this point of the target sentence results in after reconsidering the preceding underspecified ignorance utterance of the relevant speaker. In other words, the comparative modifier triggers an undetectable specific ignorance implicature weaker than the one of the superlative modifier (possibly also of a different nature), which might cause a detectable infelicity at a different region compared to the superlative case due to reconsideration of the underspecified ignorance context. Whenever readers re-read the PP they slowdown in the comparative items of the ignorance condition, because only then do they calculate (off-line) the interpretation of the modified numeral with the specific ignorance signal and notice the infelicity of the specificity of the derived interpretation given the prior underspecified

ignorance utterance.

Therefore, whether we take the observed effects in Region 8 to be the on-line manifestation of the derivation of the specific ignorance implicature (H3) or an indirect manifestation of such a derivation because of the caused infelicity of the resulting interpretation of the target sentence given the preceding context sentence (H2), we can conclude the following: They are to be associated with the derivation of a specific speaker ignorance implicature (HH2); one that is weaker (I1) and/or of a different nature (I2) compared to the specific ignorance implicature of the superlative modifier.

Finally, if we associate the processing penalty under discussion with the derivation of a specific speaker ignorance implicature that is weaker and/or different in nature than the specific ignorance implicature of superlatives, our findings in question are obviously in line with Ciardelli et al.'s (2017) account. Moreover, they go against all those accounts that maintain that comparative modifiers do not give rise to speaker ignorance implications (e.g., Büring, 2008; Coppock & Brochhagen, 2013b; Cummins & Katsos, 2010; Geurts & Nouwen, 2007; Kennedy, 2015; Nouwen, 2010, see section 2.2). There is yet another alternative explanation, however, that is in some sense more favorable to all these accounts.

Alternative explanation: Effects in Region 8 due to Manner implicature. The slowdown we find in late measures in Region 8 at the expense of the comparative condition might not be connected with the derivation of an ignorance implicature but rather with a different type of reasoning comprehenders get engaged in. It could very well be that the effects occurring after the critical regions (6 and 7) have been read reflect that comprehenders had an afterthought: "*why did the speaker not use 'minstens' (at least), which is a better cue to speaker ignorance?*". Such a reasoning results in a Manner implicature, apparently interfering with on-line processing. Hence, there is some competition going on between the two numeral modifiers, with the comparative modifier (*meer dan*) being pragmatically less appropriate in ignorance contexts, either because the ignorance implicature it triggers is weaker or because it triggers no ignorance implicature at all contrary to what the context prescribes.

If we take the comparative modifier to trigger no ignorance implicature, our findings seem to be compatible with the aforementioned theories predicting no ignorance implication with comparative modifiers (e.g., Büring, 2008; Coppock & Brochhagen, 2013b; Cummins & Katsos, 2010; Geurts & Nouwen, 2007; Kennedy, 2015; Nouwen, 2010). However, this interpretation is not quite appealing as it is in direct contradiction with Westera & Brasoveanu's (2014) off-line findings as to ignorance interpretations of comparative modifiers.

Assuming, on the other hand, that *more than* gives rise to ignorance implications, only less robustly, and thus less appropriately, than *at least*, is again in favor of Ciardelli et al.'s (2017) account. A similar competition between expressions that are less good and better cues to a certain implicature has been

brought up in the experimental pragmatics literature before (e.g., Degen & Tanenhaus, 2011, 2014). In particular, the competition in question brings to mind the contrast in implicature rate found by Degen & Tanenhaus for the scalar expressions *some* and *summa*, with the latter obtaining more *not all*-implicature responses (not associated with longer response times, though).

We realize that the ignorance implicature associated with the comparative modifier has possibly been triggered too, prior to the Manner-based reasoning of comprehenders, but was apparently too weak to be traced on-line (cf. negative interaction and no Context effect for comparative NM in previous region). Perhaps this explains why the reasoning in question was manifested (in late measures and) in Region 8.

After discussing in detail each effect of interest attested in Regions 7 and 8 and considering a number of possible explanations, it seems that a joint treatment of these effects speaks in favor of hypothesis HH2, while hypothesis HH1 leaves the effects in Region 8 unexplained.

Region 7: Positive main effect of NM. The last effect to consider, which is of less importance as it does not relate to the hypotheses we are interested in in the present experiment, is the positive effect of the NM factor attested in the probability of re-reading Region 7. Notice the big difference between the two NM conditions in the baseline authority Context condition at the phrase *ten people* in Figure 6.2 (in gray). This indicates that readers were more likely to re-read *ten people*—the region with the numeral phrase, where the interpretation of the whole modified numeral is completed—in comparative items than in superlative items. I should note that such an effect is unexpected under accounts that derive obligatory and robust ignorance effects with superlative modifiers and no or less robust ignorance effects with comparative modifiers (Ciardelli et al., 2017; Coppock & Brochhagen, 2013b; Geurts & Nouwen, 2007; Nouwen, 2010). If anything, we would expect the relevant difference to be in the opposite direction. But what could have caused this difference, given that the meaning of a comparative target sentence and of a superlative target sentence is equivalent in the authority condition? We argue that the effect in question is to be associated with the fact that, all else being equal, *minstens* ('at least'), but not *meer dan* ('more than'), states and does not exclude the minimum value *exactly* n of the relevant range (i.e., the minimum value compatible with the modified numeral). This might seem more appropriate, since the superlative modifier draws attention to a number that would make the statement true, unlike the comparative modifier.

One could perhaps take it that the positive NM effect partly drives the attested negative interaction effect (in the same measure and region), assuming that the factor responsible for the NM effect is other than what we argued above, that is, it is specific to the authority Context condition and does not influence the ignorance condition. Given this possibility, we would not be able to draw any firm conclusions as to the differences we observe, discussed in the

foregoing, as the baseline for the comparative condition would be faulty. However, there is no obvious, convincing factor that would make us consider such a scenario.

6.8 Recap and overall conclusions

Experiment 2 was conducted to test whether the Context effect attested in experiment 1 is to be associated with the derivation of a specific speaker ignorance implicature. In order to do so, we also tested the comparative counterpart of the Dutch lower-bound superlative modifier tested in experiment 1. We assumed that if the attested Context effect was due to the infelicity of the specific core meaning of the superlative modified numeral given the preceding underspecified ignorance utterance of the speaker, the two NM conditions should behave alike. This is so because the core meaning of the comparative modifier is as specific. Moreover, if the observed Context effect had been due to the infelicity of the specific ignorance implicature triggered by the target sentence given the preceding ignorance context sentence or due to the on-line computation of such an implicature, the two NM conditions should behave differently. This is the case because that implicature is argued to be triggered less robustly with comparative than with superlative modifiers. We further got rid of other possible confounds of the Context effect in question by making sure that the number to be modified by the numeral modifier in the target sentence is a round and imprecise number, and that the relevant quantity expressed in the target sentence is in some comparison with a quantity given in the Intro of the texts, making its mention more expected and natural given a speaker ignorance context.

Experiment 2 revealed a different processing profile between the comparative and the superlative conditions at the same measure (probability of re-reading) and region (numeral phrase) where the Context effect of experiment 1 occurred: The difference between ignorance and authority items was greater in the superlative condition than in the comparative condition. Also, readers were more likely to re-read the numeral phrase (Region 7) modified by *at least* in the ignorance than in the authority condition, as attested in experiment 1, too. These findings are taken to be an indication of the derivation of a non-obligatory, context-dependent specific speaker ignorance implicature at the numeral phrase modified by *minstens* in a speaker ignorance context—manifested either on-line or indirectly through the infelicity of the resulting interpretation of the target sentence given the preceding ignorance utterance. Once more, we find evidence that speaker ignorance effects are pragmatic in nature.

Interestingly, the difference in the processing profile of *more than* and *at least* does not stop here. The reverse processing behavior was attested in Region 8 between the superlative and the comparative conditions. Readers slowed down when re-reading the PP in Region 8 in an ignorance as opposed to an authority context and this slowdown was bigger in the comparative than in the superlative items. We interpreted these findings as indicating that another

pragmatic process is taking place at a later point of the target sentence at the expense of comparative items. Either a weaker ignorance implicature and/or of a different nature is generated with the comparative modifier as compared to the superlative modifier (manifested on-line or indirectly, see explanation for *at least* above), or readers, on second thought, figure after the encounter of the comparative modifier in an ignorance context that the superlative modifier would have been a better cue to ignorance (encouraged by the relevant context), thereby generating on-line a Manner implicature. We further speculated that the comparative modifier provides a less appropriate cue to a speaker ignorance implicature either because the ignorance implicature it triggers is less robust compared to that of the superlative modifier or not available to begin with, as Büring (2008); Coppock & Brochhagen (2013b); Cummins & Katsos (2010); Geurts & Nouwen (2007); Kennedy (2015); Nouwen (2010) have claimed.

Putting all our findings together, it turns out that we need a **composite theory** of ignorance inferences of superlative and comparative modifiers. Such a theory should account for the following findings of the present study:

(i) Superlative and comparative numeral modifiers trigger specific speaker ignorance implications

(ii) Comparative numeral modifiers trigger weaker/different specific ignorance implications than those of superlative modifiers[6]

(iii) Specific ignorance effects of numeral modifiers are non-obligatory, context-dependent inferences

How can we capture (i)-(iii) taking into consideration the set of the theoretical accounts we have been discussing so far? As far as (i) is concerned, that superlative modifiers trigger specific speaker ignorance is in favor of Büring (2008); Ciardelli et al. (2017); Geurts & Nouwen (2007); Kennedy (2015); Nouwen (2010, 2015); Schwarz (2016a); Spector (2015); Spychalska (2015), but against Cohen & Krifka (2014); Coppock & Brochhagen (2013b), while the finding that comparative modifiers trigger similar speaker ignorance inferences is only compatible with Ciardelli et al.'s (2017) account. So is the finding in (ii). That is, Ciardelli et al.'s (2017) account is the only one to capture the strength difference (and the difference in nature) of ignorance implicatures between superlative and comparative modifiers. Nevertheless, (iii) goes against Ciardelli et al. (2017); Cohen & Krifka (2014); Coppock & Brochhagen (2013b); Geurts & Nouwen (2007); Nouwen (2010); Spector (2015); Spychalska (2015), whereas it is in line with Büring (2008); Kennedy (2015); Nouwen (2015), and Schwarz (2016a).

In this context, I propose that the composite theory should incorporate an analysis à la Nouwen (2015) for superlative modifiers, and a standard Quantity-based analysis for comparative modifiers like Büring's (2008), Kennedy's (2015),

[6]This is the case considering either the ignorance implicature explanation or the Manner implicature explanation for the relevant effects in Region 8.

and Schwarz's (2016a) proposals for superlative modifiers (which are basically equivalent to the comparative part of Ciardelli et al.'s (2017) story). Let me explain why this is so. First of all, these accounts do derive a specific ignorance implicature, so (i) is covered. More importantly, this particular combination of accounts also covers (ii) and (iii), that is, as regards the strength of ignorance with each numeral modifier separately but also comparatively. To elaborate on that, apart from its standard Quantity-based pragmatic component, Nouwen's (2015) proposal for superlative modifiers primarily encompasses a somewhat conventional basis of speaker ignorance implicatures, namely, the (Manner-like) pragmatic reasoning about the presupposition-based alternatives. As I already noted in section 5.3.2.1, Nouwen fails to be explicit about how his anti-specificity presupposition relates to the Manner-like reasoning and in particular what this means for the strength of the resulting ignorance implicature. Based on experimental evidence on Manner implicatures, I took the Manner-like reasoning in Nouwen's (2015) account to be non-obligatory and context-dependent. Now, I moreover argue that the presupposition-based component of the entire ignorance implicature is responsible for the resulting strength of ignorance implicatures of superlative modifiers as compared to comparative modifiers. The idea is that speaker ignorance is more robust with superlative rather than with comparative modifiers, or in other words, context-permitting, it is more likely to draw a presupposition-based Manner-like reasoning than a standard Quantity-based reasoning (i.e., with Horn-scale alternatives), because, contrary to the latter case, in the former case there is a linguistic signal encouraging the relevant reasoning.

Pursuing the composite theory, we can further account for a number of other empirical observations and findings that have been mentioned earlier. First, we can explain the putative difference between the two numeral modifiers as concerns the availability of the exhaustive interpretation of the minimum value of the range (of values) expressed as necessary part of and responsible for their specific speaker ignorance interpretation. This stems from our findings in chapter 3 (remember motivation of eye-tracking experiment 2 in section 6.2.1 under CONF3 paragraph): I.e., (i) that the minimum value of the relevant range is a necessary alternative fed into the pragmatic mechanism those numeral modifiers involve, and (ii) that the exhaustive interpretation of this value being a necessarily available (deontic or witness) possibility is more robust with superlative than with comparative modifiers as far as their so-called (modal) variation or scalar implicatures are concerned. These findings provide insights into the nature of the pragmatic mechanism numeral modifiers involve, which, crucially, depending on the presence of other operators and on what operators those are, outputs different implications. Given that, we assume that the contrast under consideration carries over into the case of ignorance implications (see similar claim in Ciardelli et al., 2017).

Let us see now how the composite theory accounts for the difference. Take the target sentence of the example item in experiment 2, slightly modified and repeated in (9) below.

(9) Wesley tattooed at least 10/more than 9 people.

The relevant alternative propositions for (9) are the following (where $[n]$ stands for the proposition with *exactly n* and $[n, ...)$ for the proposition with *at least/more than n*):

(10)

$$\left\langle \begin{array}{ccccc} \ldots & [12] & [11] & [10] & \\ \ldots & [13,\ldots) & [12,\ldots) & [11,\ldots) & [10,\ldots) \end{array} \right\rangle$$

After applying the standard recipe to the alternatives, we derive primary Quantity implicatures for each of the alternatives and specifically arrive at the implications $\neg\Box_{Bel}[10]$ and $\neg\Box_{Bel}[11,\ldots)$, which combined with the Quality implication $\Box_{Bel}[10,\ldots)$ gives $\Diamond_{Bel}[11,\ldots)$ and $\Diamond_{Bel}[10]$, the epistemic possibility as to the minimum value of the relevant range. I.e., the latter is a defining part of the resulting (and familiar until now) specific ignorance reading. Crucially, this inference is predicted to be stronger in the case of superlative modifiers on Nouwen's (2015) account than in the comparative modifiers on a standard Gricean account for superlative modifiers (as in Schwarz, 2016a) by virtue of the following difference: The alternative in (one of) the relevant premise(s) $\neg\Box_{Bel}[10]$ in the case of superlative modifiers comes from a scale that is more closely tied to the lexical meaning of the superlative modifiers, as it becomes available via a(n anti-specificity) presupposition, while in the case of comparative modifiers it is part of the Horn scale < *exactly*, *more than* >. As previously noted, given this presupposition there is an extra linguistic signal in the case of superlatives encouraging the reasoning about *exact* alternatives, hence, $\neg\Box_{Bel}[10]$, a necessary constituent for the resulting specific speaker ignorance implication, $\Diamond[10]$ and $\Diamond[11,\ldots)$. Therefore, context permitting (cf., e.g., authority vs. ignorance context), the inference $\Diamond[10]$ is predicted to be more readily available in the case of superlative modifiers as compared to comparative modifiers by virtue of having a somewhat conventional basis.

In general, by adopting the composite theory, we can account for the cancellability resistance of any implicature of the form $\neg\Box_{Bel}[n]$ of superlative modifiers (see (11-a)) as contrasted with the cancellability (tolerance) of the implicatures triggered by alternatives associated with a Horn-scale (as in (11-b)–(11-c)), either in comparison with comparative modifiers or in comparison with superlatives' implicatures of the form $\neg\Box_{Bel}[n, ...)$.

(11) a. At least 50 people came to the party. ??Actually, there were 53 people at the party.
b. More than 50 people came to the party. Actually, there were 53 people at the party.
c. At least 50 people came to the party. Actually, there were at least 53 people at the party.
d. More than 50 people came to the party. Actually, there were (even)

more than 53 people at the party.

The fact that the continuation sentence seems to be infelicitous to some extent only in (11-a) suggests that only the implicature $\neg\Box_{Bel}[53]$ with superlative modifiers is a robust implicature as compared to the implicatures $\neg\Box_{Bel}[53]$ with comparatives, $\neg\Box_{Bel}[53, ...)$ and $\neg\Box_{Bel}[54, ...)$, respectively for the rest of the cases. As already mentioned, Nouwen (2015) makes a similar point as to the defeasibility of implicatures triggered by Horn-numeral-scale alternatives vs. implicatures triggered by (anti-specificity) presupposition-based alternatives of superlative modifiers. Note also that—apart from deriving a context-independent specific ignorance implicature—Ciardelli et al. (2017), Spector (2015) and Spychalska (2015) fail to capture the difference between (11-a) and (11-c), as the disjunctive alternatives they posit for an *at least* utterance come from the same source.

To conclude, the present chapter adequately answers the following questions from the Introduction (chapter 1) as to speaker ignorance effects:

- ▷ How are ignorance inferences accessed in incremental interpretation of modified numerals? What are the insights into the nature and strength of speaker ignorance inferences and into the underlying mechanism of derivation? How is this mechanism different across different modified numerals? How does the mechanism responsible for speaker ignorance effects relate to that responsible for variation effects?

Our overall results point to the availability of specific speaker ignorance effects. They further shed light on the nature of such ignorance inferences and the pragmatic mechanism numeral modifiers involve. They reveal that this mechanism is not obligatorily put into operation with every use of a superlative modifier and suggest that it is more loosely associated with comparative numeral modifiers than with superlative numeral modifiers. More generally, our experimental investigation of ignorance implications brings to light (direct or indirect) traces of pragmatic reasoning in incremental interpretation and our findings suggest that pragmatic processes can be attested on-line or indirectly: Possibly in the form of a non-obligatory Quantity-based specific ignorance implicature with superlative modifiers, or as a weak(er) ignorance implicature with comparative modifiers, perhaps also of a different nature to some extent, or as a Manner-based reasoning, implying a less robust (or, less likely, not available) ignorance implicature. In order to account for all these findings, embracing the relevant indications, I put forth the so-called composite theory, which basically outlines the pragmatic mechanism each type of numeral modifier is associated with. Although the pragmatic mechanism the numeral modifiers we tested involve is similar and employs the same set of alternatives, it differs in the source of part of these alternatives (i.e., of the *exact* alternatives). Depending on whether numeral modifiers appear in embedding environments (e.g., universally quantified ones) or non-embedding environments, the relevant pragmatic mechanism outputs specific partial variation and specific speaker ignorance implicatures,

respectively. Thus, the composite theory also accounts for the strength difference between superlative and comparative modifiers in specific partial variation effects given the results reported in chapter 3.

Finally, our on-line findings come to supplement studies that draw similar conclusions with respect to accessing various types of pragmatic reasoning in on-line processing (e.g., Breheny et al., 2006; Engelhardt, Bailey & Ferreira, 2006; Fukumura & van Gompel, 2017; Huang & Snedeker, 2009; Panizza et al., 2009, a.o., on other types of Quantity-based reasoning). Crucially, our findings further demonstrate that pragmatic reasoning manifests itself late in the eye movement record, i.e., in late measures such as the probability of re-reading and re-reading times.

This chapter came to complete our investigation of speaker ignorance inferences, after having also put forth the composite theory to account for them as well as for the variation effects triggered by superlative and comparative numeral modifiers. We will next turn to a third type of inference that seems to be available with numeral modifiers, the so-called speaker indifference effects. The following chapter is concerned with these unexplored effects in the domain of modified numerals and sets out to investigate them both theoretically, by extending the composite theory from the current chapter, and experimentally.

CHAPTER 7

Speaker indifference effects of numeral modifiers

7.1 Introduction

As already pointed out, the semantico-pragmatic literature on modified numerals has mostly been dealing with accounting for (i) the availability of speaker ignorance effects with superlative modifiers (Geurts & Nouwen, 2007; Coppock & Brochhagen, 2013b; Cummins & Katsos, 2010; Nouwen, 2010), see first implication in (1), (ii) the absence of scalar implicatures with unembedded occurrences of modified numerals (Krifka, 1999; Fox & Hackl, 2006), see second implication (1), and (iii) their reoccurrence when modified numerals appear in certain embedding environments like that in (2) (Büring, 2008; Fox & Hackl, 2006; Mayr, 2013; Nouwen, 2015).

(1) Trump won at least 270 electoral votes.

$\rightsquigarrow$ *The speaker doesn't know the exact number n of electoral votes Trump won*
$\not\rightsquigarrow$ *Trump did not win at least 271 electoral votes*
(+ (1) $\rightarrow$ *Trump won exactly 270 electoral votes*)

(2) Trump has/needs to win at least 270 electoral votes (in order to become the new US president).

$\rightsquigarrow$ *Trump does not have to win more than 270 electoral votes*

Interestingly, the relevant literature has overlooked an additional type of inference that seems to be available with modified numerals, which, similarly to speaker ignorance, exhibits a modal flavor. This novel inference is what I will

call *speaker indifference*. To illustrate, (1) can also convey that all that is relevant according to the speaker or all that the speaker cares about is that $n \geq 270$ and the speaker does not care about the exact number of electoral votes Trump won. Considering the context of the latest American presidential election and knowing that *it will take 270 electoral votes to win the 2016 presidential election*[1], on such an interpretation (1) signals that Trump is the new president of the US. This inference is distinct from the relevant speaker ignorance inference, as demonstrated by the discourse below:

(3) Early this morning, Donald Trump won at least 270 Electoral College votes, giving him the majority he needed to become president of the United States.[2]

Crucially, the speaker in (3) is probably informed about the exact number of electoral votes Trump won (cf. felicitous follow up in (4)), or at least she appears to know the current number of counted votes for Trump.

(4) Actually, the final Electoral College totals are 232 for Hillary Clinton and 306 for Trump.[3]

It is not surprising that this type of modal inference is attested in the domain of modified numerals. Items that exhibit speaker ignorance and free-choice(-like) effects usually also exhibit speaker indifference effects. Examples of such items are the free relatives with *-ever*, free-choice items, and disjunction (see Aloni, 2007 for unembedded disjunction, Lauer, 2009 for free relatives with *-ever*, and Aloni & van Rooij, 2007; Chierchia, 2006; Choi, 2005; Kratzer & Shimoyama, 2002 for existential free choice indefinites). Given that superlative numeral modifiers like *at least* have already been noted to exhibit the aforementioned family of modal inferences, i.e., speaker ignorance (Büring, 2008; Geurts & Nouwen, 2007) and free choice(-like) effects (Buccola & Haida, 2017; Nouwen, 2015, see also chapter 3), one would expect that they could also trigger speaker indifference implications.

In the rest of this introduction (sections 7.1.1 and 7.1.2) and before moving to the core of the chapter, I will attempt to accurately define the notion of indifference we will be concerned with.

7.1.1 Speaker indifference vs. agent indifference

Speaker indifference has not been much discussed in the semantics/pragmatics literature in general. If you think otherwise, then what you probably have in mind is the so-called *agent indifference*. The majority of the relevant literature—with the exception of some mentions or in-depth studies in the few works cited

[1] Adapted from http://www.270towin.com.

[2] Example taken from http://www.hopeferdowsian.com/how-do-we-move-forward/.

[3] Adapted from https://www.businessinsider.nl/final-electoral-college-map-trump-clinton-2016-11/.

in the previous section—deals with agent indifference. In fact, the aforementioned domains of free choice indefinites and free relatives happen to exhibit both types of indifference (for agent indifference with epistemic or free choice indefinites see Alonso-Ovalle & Menéndez-Benito, 2011; Chierchia, 2013; Choi, 2007, and with free relatives see von Fintel, 2000; Condoravdi, 2015, for instance). The current section serves to point out that the speaker indifference inferences we are interested in with respect to the domain of modified numerals are not to be confused with the more familiar agent indifference inferences from other empirical domains. In doing so, I do not intend to make any claim about whether or not agent-oriented and speaker-oriented indifference should be derived from the same mechanism in the empirical domains that exhibit both types of indifference.

Let us take a look at an example of an agent indifference implication. Below I exemplify with a free relative with -*ever* from Condoravdi (2015):

(5) Ed just voted for whoever was at the bottom of the list.

$\rightsquigarrow$ *Ed voted indiscriminately for the person at the bottom of the list*

By this example, Condoravdi (2015) shows that the inference drawn in (5) has to do with the agent of the sentence, Ed, as opposed to the speaker: It is Ed who was indifferent to or did not care about the identity of the person that was at the bottom of the list, and not the speaker of the utterance. Notice that this type of indifference does not seem to arise with *at least*, in a similar example to (5). The relevant agent indifference reading would be that *Stephanos was indifferent to the exact number of books he grabbed.*

(6) Stephanos grabbed at least two books from the pile.

$\not\rightsquigarrow$ *Stephanos doesn't care how many books he ended up getting, as long as those were 2 or more*

As is obvious, such a reading does not become available in (6) Hence, we can only have indifference on the part of the speaker with *at least*, i.e., speaker indifference.

In general, authors studying indifference in the various empirical domains focus on one type of indifference: Those who analyze agent-oriented indifference do not talk about speaker indifference, and vice versa. Choi (2005) explicitly makes a distinction between the two types of indifference and might be the only one that studies both of them. Alonso-Ovalle & Menéndez-Benito (2011), too, make a deliberate distinction between the agent indifference reading of Spanish indefinite *un* NP *qualquiera* and an additional reading, which is speaker-oriented and they characterize it as the *unremarkable reading.* Example (7) from Alonso-Ovalle & Menéndez-Benito (2011) illustrates both readings of the Spanish indefinite in question:

(7) Juan cogió un libro cualquiera.
Juan grabbed a book CUALQUIERA

$\rightsquigarrow$ *Juan took a random book* (agent indifference reading)
$\rightsquigarrow$ *Juan took a book, which was not special or remarkable according to the speaker* (unremarkable reading)

Although Alonso-Ovalle & Menéndez-Benito dismiss the latter implication and suggest that it could actually be a *depreciative* reading in Haspelmath's (1997) classification, it is rather clear that this reading has something in common with speaker indifference. Namely, both readings share the fact that it is the speaker who is taken to consider the precise identification of the entity or the quantity in question as uninteresting or irrelevant. This will be the type of indifference we will be concerned with in the present chapter.

7.1.2 Speaker indifference vs. speaker uncooperativity

Having clarified that speaker indifference is the type of indifference modified numerals exhibit, I would like to make an attempt to better define that speaker-oriented reading. More precisely, in this section, I aim to discuss how speaker indifference is similar to or different from other related implications about the speaker's stance, such as speaker's *uncooperativity* or *uncooperativeness* (Fox, 2014; Simons, 2005; Starr, 2016) or *unwillingnesss to inform* (Condoravdi, 2015; Lauer, 2013).

Theorists have identified cases like (8), for instance, which are not taken to signal ignorance on the part of the speaker (which is the most usual interpretation of unembedded disjunction), when the speaker is a quizmaster.

(8) There is money in box 20 or 25. (Fox, 2014)

$\nrightarrow$ S *doesn't know whether there is money in box 20 or box 25*

The speaker in (8) seems to violate Grice's Cooperative Principle (CP), expected to be observed by all participants in a conversation. More specifically, as Fox (2014) points out, (8) deactivates or cancels the first Maxim of Quantity, as the competent speaker does not reveal all the relevant information she has available. Actually, Grice (1975) himself remarks that the speaker *may opt out from the operation both of the maxim and of the CP; [...] is unwilling to cooperate in the way the maxim requires [...] may say, for example, I cannot say more; my lips are sealed* (p. 49). The ignorance implication in (8) does not go through exactly because the maxim it is based on is disabled.

But is the CP disabled in the case of an indifferent speaker too? Take (3) again, repeated below as (9). The speaker wants to convey that Trump will be the next US president. Given that, she is being cooperative and willing to share the information that is necessary (and sufficient) for making such a conclusion, and this is that (already in the morning of November 9, 2016) the number of electoral votes Trump had won was ≥ 270.

(9) Early this morning, Donald Trump won at least 270 Electoral College votes, giving him the majority he needed to become president of the United States.

On the contrary, a speaker, who is a quizmaster and utters (10), is not revealing all the information she apparently knows (and is needed for winning the game), because being cooperative and sharing all the necessary information will ruin (the end of) the game.

(10) There is money in at least two boxes.

Crucially, there is a rather sharp difference in strength between speaker indifference, on the one hand, and speaker uncooperativeness or unwillingness to reveal information, on the other, in the sense that in the latter case the speaker appears to be more reserved towards the other members of the discourse compared to the former case. That is, while speaker indifference is about what the speaker does not consider to be relevant or necessary information, i.e., that she considers that she does *not need* to mention the precise information, speaker uncooperativeness has to do with the speaker considering that she *needs* to *not* mention the precise information. Hence, speaker indifference and speaker uncooperativity/unwillingness to inform are distinct and we shall not analyze speaker indifference readings as uncooperative utterances.[4]

In the following, I sketch a tentative analysis of speaker indifference implications of modified numerals (section 7.2), by extending the composite theory proposed at the end of the previous chapter (section 6.8). Section 7.3 presents the results of our eye-tracking experiment 2 concerning the *indifference* Context condition, whose inclusion aimed at tracking down speaker indifference effects in the incremental interpretation of numeral modifiers. In section 7.4, I conclude the present chapter by bringing together what has been discussed in this chapter and revealed by experiment 2, and I also bring up further points that seem to relate to speaker indifference.

7.2 An analysis of speaker indifference: Extended composite theory

In this section, to analyze speaker indifference implications of modified numerals, I will tentatively extend the composite theory proposed in section 6.8 to capture speaker ignorance implications with both superlative and comparative modifiers, which also accounts for their partial variation effects in embedding contexts. The idea is that all these three types of inferences are derived via the same pragmatic mechanism. Adopting such an idea, apart from making up

[4]Interestingly, cases of speaker uncooperativity and unwillingness to inform have been argued to involve pragmatic reasoning and have been analyzed by means of pragmatic frameworks that do not appeal to CP and subordinate maxims, cf., for instance, optimization-based accounts in Lauer (2013, 2014), Asher (2012), and Asher & Lascarides (2013).

an elegant and neat story, is in line with, and actually somehow suggested by, a number of other theories accounting in a unified manner for the availability of the corresponding inferences in other empirical domains. For instance, Kratzer & Shimoyama (2002), Chierchia (2006) as well as Aloni & van Rooij (2007) propose the use of the same machinery to capture free choice effects, speaker ignorance and speaker indifference effects in the domain of existential free choice items. As will be noted below, the analysis in the present section will have some aspects in common with these accounts.

First, I take the (unembedded) uses of *at least* by a cooperative speaker to signal different types of effects or different reasons why she did not mention an exact number instead. It might be so because she does not consider it to be true or relevant.[5] This will be at the core of the present analysis (in a similar way to Aloni & van Rooij, 2007), aiming to capture those two reasons in a unified way. Importantly, speaker indifference, as formulated here but also in other works (cf. Aloni, 2007 on unembedded disjunction; Lauer, 2009 on free relatives with *-ever*; Kratzer & Shimoyama, 2002; Lauer, 2012 on existential free choice items), involves and alludes to the notion of relevance. That is, speaker indifference about a specific x (be it a referent or a cardinality) is taken to mean that the speaker considers that it is not relevant to specify x. The notion of relevance will play a key role in the present analysis.

Moreover, together with Kratzer & Shimoyama (2002) and Aloni (2007) among others, I assume that speaker indifference has a pragmatic status. This has been convincingly argued and experimentally corroborated for speaker ignorance effects of numeral modifiers (see chapters 4, 5, and 6). The availability and the pragmatic nature of speaker indifference effects are evidenced by its disappearance in downward entailing environments, its cancellability, and its reinforceability, to be demonstrated in the following for *at least*. In (11) we embed our toy example (1) in the antecedent of a conditional. The main interpretation of (11) seems to be one where the speaker states the requirement for US presidency (with respect to the number of electoral votes—irrespective of her own beliefs and preferences), here applied to Trump. More accurately, assuming a universal quantification analysis of conditionals (see also Mayr, 2013), (11) conveys that *in all worlds where the number n of electoral votes Trump wins is* ≥ 270*, he is the new US president.* So as is obvious, embedding sentence (1) in (11) makes the speaker indifference reading disappear or weaken severely.

(11) If Trump wins at least 270 electoral votes, he will become the new US president.
$\rightsquigarrow$ *in all worlds where $n \geq 270$* AND *the speaker does not consider n is*

[5]This comes very close to what Lauer (2013, 2014) dubs *Need a Reason* (NaR) implicatures of unembedded disjunction. In his discussion of unembedded disjunction uses, he posits that a speaker needs to have a reason for choosing such an expression over the alternative disjuncts, and possible reasons are speaker ignorance, speaker unwillingness to reveal all her information or full and precise knowledge, or even speaker's attempt to acknowledge a disagreement (known as *acknowledging a disagreement* or *agree to disagree* cases).

relevant, Trump is the new US president

Moreover, the speaker indifference reading of (1) seems to be cancellable by a continuation like that in B's utterance in (12), where B cancels the inference that she doesn't consider $n = 300$ to be relevant (or that $n > 270$ is relevant) by stating the exact opposite and without sounding self-contradictory. Note that B gives the impression that she is knowledgeable about the exact number of electoral votes Trump won, so an ignorance inference is less likely to arise here. This is further accomplished by taking B's utterance as an answer to a polar question, which has been argued to be an environment that weakens or makes speaker ignorance interpretations of numeral modifiers unavailable (Westera & Brasoveanu, 2014).

(12) A: Did Trump win the election?
B: Yeah, he did win at least 270 electoral votes. In fact, it's also significant that he exceeded the 300 electoral votes!

Finally, the speaker indifference reading can also be reinforced, see (13), by stating explicitly what the speaker does not consider to be important/relevant or does not care about in the current purposes of the exchange.

(13) Trump won at least 270 electoral votes. This is all that matters! The exact number of votes he won is not that important, it makes no change whatsoever. (In any case, I can tell you how many he won if you want to know the exact number.)

In the foregoing, I have argued and shown that speaker indifference effects are pragmatic inferences available with *at least* in non-embedding contexts (where speaker ignorance inferences have been suppressed), and that the indifferent speaker, while being cooperative, does not specify the exact quantity under discussion because she thinks it is not relevant. In the following, I will extend the composite analysis from the previous chapter on ignorance so that it can also capture speaker indifference. That is, I will build on the neo-Gricean accounts of speaker ignorance as Quantity implicatures, which proved to fare best in the experiments presented in chapters 5 and 6.

Drawing on Quantity-based accounts of speaker ignorance implications of *at least* such as Nouwen's (2015) or Schwarz's (2016a), I propose to derive speaker indifference effects as primary Quantity implicatures. The alternatives to an *at least n* utterance that the speaker chose not to make result from the interaction of the Horn number scale in (14) and the scale in (15).

(14) {one, two, three, ...} (natural number scale)

(15) {at least, only}

The resulting alternative propositions from the interaction of the previous scales, ordered in terms of strength/informativity, are:

(16)

$$\left\langle \begin{array}{lll} \ldots\ [n+2] & [n+1] & [n] \\ \ldots\ [\text{n+3, ...}) & [\text{n+2, ...}) & [\text{n+1, ...}) \end{array} \quad [\text{n, ...}) \right\rangle$$

($[n]$ stands for the proposition with *only* n and $[n, ...)$ stands for the proposition with *at least* n). Given that, the subsequent Quantity-based reasoning might be drawn by a listener:

(i) The speaker believes $[n, ...)$ to be true, i.e., $\Box_{Bel}[n, ...)$ (Quality)

(ii) The speaker did not make a stronger claim uttering one of the stronger propositons in (16), because it is not the case that she believes them to be true. That is, $\neg\Box_{Bel}[n]$, $\neg\Box_{Bel}[n{+}1]$, ..., $\neg\Box_{Bel}[n{+}1, ...)$, $\neg\Box_{Bel}[n{+}2, ...)$, ... (primary Quantity implications)

Hence, the listener arrives at the conclusion that the speaker does not believe that $[n]$ is true and she doesn't believe that $[n+1, ...)$ is true (note that the primary implication for each alternative stronger than $[n+1, ...)$ in (16) is already entailed by the primary implication $\neg\Box_{Bel}[n+1, ...)$).

At this point, I am going to add a second option/reason to the Quantity assumption the listener makes in (ii), namely, that the speaker did not make a stronger claim because she does not consider the stronger propositions in (16) to be relevant. This comes from the Gricean maxim of Relation, instructing interlocutors to *be relevant*; here the listener's relevant assumption is that the speaker considers that $[n, ...)$ is relevant. In order to capture both the belief and the relevance restriction/criterion of the Quantity maxim, I propose to replace the $\Box_{Bel}$ operator that comes from Quality with a more general operator $\Box_{Assert}$. $\Box_{Assert}\ p$ stands for *the speaker believes p to be true and considers p to be relevant.* $\Box_{Assert}$ is akin to the LF assertoric operator Kratzer & Shimoyama (2002) and Chierchia (2006) devise in order to account for speaker ignorance and indifference inferences of existential free choice items (cf. also Krifka's (1999) similar pragmatic operator ASSERT).[6] I do not commit myself as to whether $\Box_{Assert}$ is an LF or a pragmatic operator. Plugging in $\Box_{Assert}$ in place of $\Box_{Bel}$ in the Quantity-based recipe for speaker ignorance effects, ceteris paribus, we arrive at the following primary Quantity implications:

(17) $\neg\Box_{Assert}[n]$, $\neg\Box_{Assert}[n{+}1]$, ..., $\neg\Box_{Assert}[n{+}1, ...)$, $\neg\Box_{Assert}[n{+}2, ...)$, ...

Therefore, the listener draws the following conclusions:

- All a speaker uttering $[n, ...)$ believes to be true and relevant is that $[n, ...)$, and

[6] Roughly speaking, $\Box_{Assert}$ also fuses together Aloni & van Rooij's (2007) sentential operators **K** and **D**, where $\mathbf{K}\phi$ expresses that the speaker knows that the proposition ϕ is true and $\mathbf{D}\phi$ expresses that it is relevant/desirable to the speaker that ϕ.

- the speaker does not believe that $[n]$ is true and does not believe that $[n+1,...)$ is true (*speaker ignorance*), and/or

- the speaker does not consider $[n]$ to be relevant and does not consider $[n+1,...)$ to be relevant (*speaker indifference*).

That is, $[n]$ and $[n+1,...)$ are not assertible because the speaker does not believe they are true and/or she does not consider them to be relevant. Hence, an utterance with an unembedded occurrence of *at least* can trigger either implication. Together with Kratzer & Shimoyama (2002) and Chierchia (2006), I assume that both inferences are available when the context of such an utterance is neutral/underspecified and does not provide a cue for/against a certain inference (e.g., a cue with respect to the speaker's information state, cf. (1) vs. (3)), thereby resolving the Quantity implicature in a certain way. For instance, if the context makes it clear that an interlocutor has precise knowledge of some exact quantity x and she makes an utterance with *at least* in relation to x, then the listener is highly unlikely to derive a speaker ignorance implicature, because that would contradict what she knows from the context with respect to the speaker's epistemic state. On the other hand, deriving a speaker indifference implicature would not result in such a contradiction, as that would just signal that the speaker does not think that specifying the exact quantity (that she is aware of) is relevant in the current context (as in (3)/(9)). In sum, unless a certain inference is encouraged by the context, both speaker ignorance and speaker indifference implications can be drawn by the listener given an utterance with an unembedded use of *at least*.

Speaker indifference inferences have not been tackled in the domain of modified numerals and their availability with unembedded uses of *at least* constitutes more evidence that the existing theoretical accounts deriving obligatory speaker ignorance implications with such configurations should be reconsidered. As extensively discussed in chapter 2, according to those accounts, e.g., Ciardelli et al. (2017); Coppock & Brochhagen (2013b); Spector (2015); Spychalska (2015), a use of *at least* is expected to be illicit in a context that makes clear (or implies as in (9)) that the speaker has knowledge of some exact quantity x. Specifically, Ciardelli et al. (2017) and Coppock & Brochhagen (2013b) posit the Quality maxims of interactive and inquisitive sincerity, respectively, which require that the speaker should not utter *at least* if she already knows how to resolve the issue an *at least*-utterance expresses; in other words, the speaker should not utter *at least* if she is in fact knowledgeable of the quantity under discussion. That is, (9), where the speaker appears to be knowledgeable, is predicted to be unacceptable. The same holds for Spector's (2015) account of *at least*, also deriving obligatory speaker ignorance. On this account, ignorance is derived as a primary Quantity implicature via the application of obligatory exhaustification above an LF belief operator. Lastly, Spychalska's (2015) assertibility condition associated with *at least* requires that the speaker entertains two possibilities in her epistemic state, so unembedded *at least* is not predicted

to be uttered by a knowledgeable speaker as in (9). In sum, on these accounts, not only is unembedded *at least* illicit in contexts where the speaker is knowledgeable, but also its licit unembedded uses give rise to one and only type of inference, i.e., speaker ignorance. Thus, the current form of these accounts cannot accommodate and is incompatible with speaker indifference inferences. On the contrary, although not captured by standard neo-Gricean Quantity-based accounts of *at least* either, such as Büring (2008), Kennedy (2015), Nouwen (2015) and Schwarz (2016a), such inferences are not at odds with them, as none of these accounts make any specification that would necessitate the derivation of an ignorance implicature with any unembedded occurrence of *at least*. For the same reason those accounts were found to be favored by the results of the experiments reported in chapters 5 and 6, indicating that speaker ignorance implications with *at least* are context-dependent, and thus, non-obligatory.

In any case, in order to also capture speaker indifference, Ciardelli et al. (2017), Coppock & Brochhagen (2013b), Spector (2015), and Spychalska (2015) should relax the restriction of speaker ignorance obligatoriness and tweak their analysis in one way or another; perhaps by supplementing their existing proposals with further conditions/maxims on the use of *at least*. Specifically, the inquisitive semantics accounts should adopt a different maxim that will be responsible for the derivation of the speaker indifference readings of *at least*, complementary to their respective sincerity maxim. Likewise for Spychalska's (2015) account. Lastly, in order to derive speaker indifference maintaining Spector's (2015) account, we should have a more flexible exhaustification operator that applies optionally and posit a more general assertoric operator in place of the belief operator Spector adopts from Meyer (2013). This is in fact similar to what is proposed in the present section and would allow for a unified account for speaker ignorance and speaker indifference effects of *at least* within a grammatical approach.[7]

Whether or not maintaining an account that assumes a similar status and mechanism of derivation for speaker ignorance and indifference inferences is on the right track is examined by eye-tracking experiment 2, which particularly included a speaker indifference Context condition (see section 7.3). Before turning to the relevant results from experiment 2, I would like to briefly discuss speaker indifference with comparative modified numerals, too, and consider the consequences of extending the composite theory from the previous chapter for such modified numerals.

The results of all our previous experiments show that comparative modified numerals are not that different from their superlative counterparts with

[7] An alternative way of deriving speaker indifference would be to tweak Spector's current exhaustification operation so as to be similar to what Aloni and van Rooij (2007) call *DD* operation, which outputs the set of worlds in which as few alternative propositions as possible are true, where the alternatives are speaker desirability states. This is recommended by Aloni & van Rooij (2007) as a way to derive speaker indifference with existential free choice indefinites. However, that would mean that, as in Aloni & van Rooij (2007), a different exhaustification operation should be assumed for each type of modal inference.

respect to pragmatic implications such as variation and speaker ignorance effects. Speculating that there is a common pragmatic mechanism responsible for the derivation of variation and ignorance inferences, which is also responsible for the derivation of speaker indifference inferences, we would expect that speaker indifference inferences arise with comparative modifiers as well.

Indeed, intuitively, (18) can imply that the speaker does not think it is relevant to say exactly how many electoral votes Trump won, as what is relevant is that Trump meets the requirement of the (minimum) number of electoral votes for becoming a US president.

(18) Trump won more than 269 electoral votes.

Also, such an interpretation seems not to arise when we embed (18) in a downward entailing context, as in (19). Instead, (19) conveys that in all worlds where the number n of electoral votes Trump gets is $n > 269$, Trump wins the presidency. The unavailability of the interpretation in question (the speaker does not care about/considers it is relevant which number in $[270, \ldots)$ n is) suggests that it is not an obligatory reading of *more than* or better that it does not have to do with the basic meaning of *more than*, but rather it is something that is pragmatically determined.

(19) If Trump gets more than 269 electoral votes — enough to cross this line — wins.[8]

$\rightsquigarrow$ *in all worlds where* $n \geq 270$ *and the speaker does not consider* n *is relevant, Trump is the new US president*

Moreover, I conducted a small off-line experiment and obtained preliminary results suggesting that speaker indifference is available with both superlative and comparative numeral modifiers. This experiment (run on Mechanical Turk with native speakers of English, see Sprouse, 2011 for collecting acceptability judgements via this platform) resembled closely the judgement task conducted by Westera & Brasoveanu (2014) to test for ignorance inferences (see section 4.2.2 in chapter 4). Participants had to read short dialogues between a person A and a person B, like that in (20):

(20) A: Did you find any marbles under the bed?

B: I found $\left\{\begin{array}{c}\text{exactly}\\ \text{more than}\\ \text{at least}\end{array}\right\}$ three marbles under the bed.

Conclusion: B thinks that it is not relevant to specify the exact number of marbles she found under the bed.

As (20) shows, the dialogue was followed by a conclusion mirroring Westera & Brasoveanu's (2014) relevant conclusion, where it constituted an ignorance inference. In our case this is an indifference inference, the indifference inference

[8]Example adapted from https://projects.fivethirtyeight.com/2016-election-forecast/.

we are after. Participants then were asked to judge how justified the conclusion is given A and B's conversation on a Likert scale from 1 to 7, where 1 is "not justified" and 7 is "strongly justified". We manipulated the numeral modifier in B's answer, testing both *more than* and *at least*, and having *exactly* as our baseline, the idea being that a speaker who uses a precisifier does consider that it is relevant to specify the exact number or even wants to emphasize the relevance of the exact number in the given context. Participants found the conclusion to be more justified when B's answer had *at least* or *more than* (means: 4.274 and 4.661, respectively) than when it had the modifier *exactly* (mean=2.032). No difference was attested between *at least* and *more than*. The former finding is taken to show that the indifference inference that the exact number is not relevant according to the speaker is an available inference of both *at least* and *more than*, as the relevant numeral modifier conditions scored significantly higher than the baseline condition, where it is clearly not the case that the exact number is not relevant according to the speaker.

The use of a polar question that weakens possible ignorance inferences of the answering utterances (cf. Westera & Brasoveanu's (2014) relevant claim and finding in section 4.2.2 in chapter 4) and perhaps also the use of perception verbs with B being the experiencer invalidate a possible criticism that the ratings we collected might just reflect speaker ignorance inferences. Note that we also included a *how many* QUD condition similarly to Westera & Brasoveanu (2014). Although we used twice as many items as they did (which were also very similar to theirs) and had a comparable number of participants (N=28), we did not find the QUD effect they did, i.e., that (highly) significantly more ignorance inferences are generated by modified numeral utterances that constitute answers to a *how many* as opposed to a polar question. Consequently, the criticism in question is untenable.

Above I argued that speaker indifference is a possible inference of comparative modifiers that is not always available, pointing to a pragmatic status. This is further supported once we apply the same diagnostics as we did with superlative modifiers. (21) and (22) demonstrate that speaker indifference is cancellable and reinforceable, respectively: The continuation sentence in (21) suggests that the speaker considers it important/worth mentioning that $n > 300$ contrary to the indifference inference where all the speaker thinks it's relevant is that $n > 269(/270)$ and that any quantity above n is not relevant/important, while the continuation in (22) highlights this very reading of speaker indifference.

(21) Trump won more than 269(/270) electoral votes. In fact, it's also significant that he exceeded the 300 electoral votes!

(22) Trump won more than 269 electoral votes. This is all that matters! Giving you the exact number of votes he won won't change the result. He's the new US president no matter what!

Although all this is already indicative of the availability and pragmatic

nature of speaker indifference effects with comparative modifiers, it would be desirable to obtain more solid evidence on this. This is additionally tested by the eye-tracking experiment 2.

Importantly, the extended composite theory predicts speaker indifference inferences to arise with comparative modifiers in a similar way to superlative modifiers. That is, they are derived via Quantity-based reasoning about stronger alternatives propositions that stem from the interaction of similar scales to those in (14) and (15), see below:

(23) $< one,\ two,\ three,\ \ldots >$

(24) $< more\ than,\ exactly >$

Assuming as part of this reasoning that a stronger alternative might not be assertible either because it is not true according to the speaker or(/and) not relevant, the listener arrives at the (primary) Quantity implications given in (17) and the conclusions below it, repeated below as (25).

(25) $\neg\Box_{Assert}[n]$, $\neg\Box_{Assert}[n+1]$, ..., $\neg\Box_{Assert}[n+1,...)$, $\neg\Box_{Assert}[n+2,...)$, ...

The listener draws the following conclusions:

- All a speaker uttering $[n,...)$ believes to be true and relevant is that $[n,...)$, and
- the speaker does not believe that $[n]$ is true and does not believe that $[n+1,...)$ is true (*speaker ignorance*), and/or
- the speaker does not consider $[n]$ to be relevant and does not consider $[n+1,...)$ to be relevant (*speaker indifference*).

Moreover, given that the composite theory has it that the corresponding scale to (24) for superlatives is tied to the lexical meaning of the modifier via a presupposition, whereas in the comparative case it constitutes a Horn scale, a strength difference between the two types of modifiers is further predicted as to the derived Quantity implications. The implication $\neg\Box_{Assert}[n]$ will be more robust with superlative than with comparative modifiers, thereby rendering the resulting specific speaker ignorance and indifference implications stronger in the former than in the latter case. No relevant difference was exhibited between these two types of modifiers in terms of strength of speaker indifference inferences by the results of our preliminary study described above, although our study of variation and ignorance inferences in the previous chapters does point to such a difference. This already suggests that the extension of the composite theory we attempt here is on shaky grounds, as such a theory predicts ignorance and indifference (and variation) to be entirely parallel. Moreover, the on-line results to be presented will turn out to be in line with the off-line data in question, as they will not reveal a strength difference between superlative

and comparative modifiers in indifference implications. That is, the attested difference with respect to ignorance implications reported in chapter 6 will not be observed for indifference implications. These results will further challenge the extended composite theory and will call for a thorough (re-)consideration of indifference inferences of superlative and of comparative modifiers, as well as of the way we tested for them in the experiment to follow.

The rest of the current chapter studies speaker indifference effects experimentally. Specifically, in the following, I will report on the results of the eye-tracking experiment 2 as regards the indifference Context condition, which were not discussed in the previous chapter.

7.3 Eye-tracking experiment 2: Indifference Context condition

As already mentioned in chapter 6, all items of the eye-tracking experiment 2 appeared in an extra Context condition, where the relevant context sentence was setting up a knowledgeable indifferent speaker. We included this additional condition in order to find out whether we can detect traces of speaker indifference effects in the incremental interpretation of numeral modifiers and investigate the timing of such traces. Moreover, we aimed to explore the possibility that speaker indifference effects have a similar status to that of speaker ignorance effects, and obtain insights into the reasoning processes and the underlying mechanisms involved in the derivation of such effects. This will enable us to assess whether it is on the right track to capture these two types of modal inferences by means of a uniform account (extended composite theory) that derives them as primary Quantity implications, importantly, resulting in the specific implicatures that the speaker doesn't have believe that $[n]$ is true and relevant, and doesn't believe that $[n+1, ...)$ is true and relevant. We further test this for both the superlative modifier and the comparative modifier, seeking to find out whether these two types of modifiers are processed and interpreted in a similar way in the appropriately manipulated environments.

7.3.1 Design

Once again, in our eye-tracking experiment 2 we had two manipulations. We manipulated:

(i) the numeral modifier in the target sentence (NM manipulation): *minstens* 'at least' vs. *meer dan* 'more than'.

(ii) the speaker's epistemic state as set up by the context before the target sentence with the numeral modifier (Context manipulation): *speaker ignorance*, *speaker indifference*, and *speaker authority* context.

Thus, the NM and Context factors were manipulated in a 2×3 design. Below I am repeating the test item (6) from section 6.2.2, presenting now all six resulting conditions.

(26) **Example item**

Intro
Wesley heeft zijn eigen zaak waar hij met veel plezier tatoeages zet. Het is er meestal erg druk en hij probeert elke dag acht mensen te tatoeëren. Deze donderdag was hij ook weer hard aan het werk.
'Wesley runs his own tattoo parlor, which he enjoys a lot. It's usually very busy and he tries to tattoo eight customers per day. Last Thursday, he was very busy.'

Context sentence

Speaker ignorance context
Ik weet niet precies hoe het met de drukte zat, maar ik heb wel een idee.
'I don't know exactly how busy it got, but I have got an impression.'

Speaker indifference context
Ik weet precies hoe het met de drukte zat, maar dat is niet zo interessant om uitvoerig na te vertellen.
'I know exactly how busy it got, but it's not so interesting to elaborate on that.'

Speaker authority context
Ik weet precies hoe het met de drukte zat en daarom zal ik je erover vertellen.
'I know exactly how busy it got, and I'll tell you about it.'

Target sentence
Wesley heeft die dag $\left\{\begin{matrix}\text{minstens}\\ \text{meer dan}\end{matrix}\right\}$ *tien mensen met veel oog voor detail*
Wesley has that day $\left\{\begin{matrix}\text{at least}\\ \text{more than}\end{matrix}\right\}$ *ten people with much eye for detail*
getatoeëerd.
tattooed.
'That day, Wesley tattooed at least/more than ten people with a real eye for detail.'

Outro
Hij heeft ook twee medewerkers in dienst die de kunst van het tatoeëeren van hem hebben geleerd.
'He also has two employees in his service that have learned the art of tattooing from him.'

The purpose of using the particular context setup is explained in section 5.3.1.

Given that, below I will focus on speaker indifference contexts and on what those are designed to achieve. First, on a par with speaker ignorance contexts that provide a cue for drawing the specific speaker ignorance implicature associated with the numeral modifier in the target sentence that follows, speaker indifference contexts bias towards the corresponding speaker indifference implicature calculation in the target sentence. Remember, speaker authority contexts provide no cue for a certain inference, but are just compatible with the core (truth-conditional) meaning of both numeral modifiers that appear in the target sentence. Also, notice that the indifferent speaker is (always) a knowledgeable speaker, as the authoritative speaker is. We set up such an indifferent speaker in order to distinguish between plain speaker indifference and speaker indifference that arises together with/because of speaker ignorance, and thus test for pure speaker indifference implications.

Additionally, the indifferent speaker always communicates that some precise quantity (of people Wesley tattooed in (26)) is not at issue, while the ignorant speaker expresses that she does not have complete knowledge of some precise quantity. The authoritative speaker always conveys that she has complete knowledge of some precise quantity, signaling it is relevant in the context, to which the ignorance and indifference implicatures favored by the other two contexts, respectively, are diametrically opposite. Hence, such implicatures are not expected to arise in the authority Context condition. This explains why authority contexts are *just* compatible with the basic meaning of the numeral modifiers in the target sentence that follows. Also, this is in line with recent experimental findings by Cremers et al. (p.c.), mentioned also earlier (see sections 5.3.1 and 5.4), revealing that both *at least* and *more than* can (highly) appropriately be used to refer to a certain quantity in scenarios where the speaker is without doubt fully knowledgeable of the precise quantity.

7.3.2 Predictions

The diagnostics we applied and the preliminary off-line data we obtained point to the availability of speaker indifference implicatures with modified numerals. The extended composite account we presented in section 7.2 specifically derives what we have called *specific implicatures* of speaker indifference, i.e., the reading *the speaker does not consider* $[n]$ *is relevant and does not consider* $[n+1,\ldots)$ *is relevant* (where n is the minimum value compatible with the modified numeral). As already specified in the previous section, given our particular context setup, a specific speaker indifference implicature will arise in the target sentence of the speaker indifference Context condition, as the context biases towards such an implicature. The extended composite theory we have considered further derives a strength difference between superlative and comparative modified numerals as to specific speaker indifference. This is predicted to be manifested as a negative interaction effect between (indifference) Context and NM at the modified numeral phrase or as a (more general) difference in the processing profile of the two NM conditions (like that observed, for instance,

in the case of ignorance).

Moreover, if speaker indifference has a similar status to speaker ignorance, involving the same implicature mechanism, then we expect the indifference Context condition to pattern with the ignorance Context condition. That is, it is predicted that, just like ignorance, the processing of indifference items will be different of that of the authority items (baseline Context condition), expecting to obtain effects similar to those reported on in section 6.6 in the previous chapter on speaker ignorance.

While considering the extended composite analysis for speaker indifference inferences (or any other unified account of ignorance and indifference inferences) we predict that indifference items should pattern similarly to ignorance items, the predictions of the theoretical accounts that derive obligatory speaker ignorance inferences with unembedded uses of *at least* are very different. Regardless of the fact that these accounts do not capture indifference inferences, we can once more test their predictions, as we now have yet another speaker knowledgeability condition. Both indifference and authority Context conditions are speaker knowledgeability conditions, that is, the speaker appears to be knowledgeable in the respective context sentences. According to the inquisitive semantics accounts (Ciardelli et al., 2017; Coppock & Brochhagen, 2013b), if a speaker utters *at least* an obligatory speaker ignorance implicature is derived regardless of Context. Given that, an incompatibility is expected to arise between the target sentence with the superlative modifier and the preceding knowledgeability Context sentence both in the authority and the indifference Context condition. No matter what processing repercussions this incompatibility has (see details in section 5.3.3.1), given these accounts we expect to see a similar processing pattern between the authority and the indifference items in the superlative condition in the critical regions. Spychalska's (2015) account predicts the same incompatibility to arise in both speaker knowledgeability conditions due to violation of the fundamental belief (assertibility) condition of the superlative modifier requiring that the speaker should entertain two possibilities in her information state. This makes the *at least* utterance illicit given a context that makes clear that the speaker is maximally informed. Hence, if anything, on this account the two knowledgeability conditions, i.e., authority and indifference, should behave the same. Rather similarly, Spector's (2015) account predicts a contradiction between the speaker's information state as set up by the authority and indifference context sentences, and that obligatorily signaled by the use of the superlative modifier in the target sentence. Thus, the respective Context conditions should exhibit a similar behavior.

7.3.3 Off-line tests

As already explained in section 6.4 of the previous chapter, by means of two questionnaires we tested how strong our manipulated contexts are in terms of speaker's knowledgeability of a certain quantity (test 1) and how coherent our constructed items are (test 2). The indifference Context was included as a third

Context condition in both questionnaires. Below, I am presenting the relevant results, which were not mentioned in section 6.4.

7.3.3.1 Test 1

In this questionnaire we tested how strong our three types of Context are in terms of knowledgeability of the quantity under discussion. Participants were given items like the test item in (26) (there was also an additional Context condition without a context sentence), where the modified numeral phrase had been replaced by the # symbol, as in (27) below.

(27) **Example item: Target sentence**
Wesley heeft die dag # mensen met veel oog voor detail getatoeëerd.
Wesley has that day # people with much eye for detail tattooed.
'That day, Wesley tattooed # people with a real eye for detail.'

They were told that # stands for a quantity and they had to indicate on the 1-7 Likert scale (1: "highly unlikely", 7: "highly likely") how likely it is that the author knows the precise quantity. Given that both indifference and authority Context conditions are knowledgeability conditions, the author of these contexts is expected to be judged as being highly likely to know the precise quantity. Thus, both conditions are expected to receive scores from the upper part of the scale and to exhibit a clear contrast with the weak knowledgeability contexts of the ignorance Context condition.

Indeed, indifference Context items received scores from the upper part of the scale (*mean* = 5.652, *median* = 6), which were highly significantly higher than those for the poor knowledgeability Context condition of ignorance ($\beta = 2.293$, $SE = .283$, $t = 8.106$, $p < .0001$). As already reported in the relevant section (6.4.1) in the previous chapter, this was found for the other knowledgeability Context condition, i.e., authority Context, as well. Interestingly, there was also found a significant difference between the two types of knowledgeability contexts, i.e., indifference and authority, such that indifference received lower scores than authority Context condition ($\beta = -.340$, $SE = .130$, $t = -2.610$, $p < .05$). From these results we conclude that indifference contexts are strong knowledgeability contexts with respect to the precise quantity mentioned, albeit somewhat less strong than authority contexts. This might have to do with the fact that the part of indifference context sentences revealing the speaker's epistemic state was not always formulated in identical words to those in the authority context sentences. Authority context sentences consisted of a part stating that the speaker has knowledge of the precise quantity and usually also of another part, a subordinate clause, explaining how she obtained this knowledge. This second clause was omitted in indifference context sentences to avoid having a very long context sentence, which should on top include a part revealing that the speaker is indifferent as to the precise quantity. Examples (28) and (29) illustrate this point (see Appendix C for more).

(28) **Authority context sentence**
Ik weet exact hoe het zit, omdat hij hierover tegen mij klaagde.
'I know exactly how it is because he complained to me about this.'
Indifference context sentence
Ik weet exact hoe het zit, maar zo interessant is dat niet om te weten.
I know exactly how it is, but it is not that interesting for you to know.

(29) **Authority context sentence**
Ik ben helemaal op de hoogte, omdat ze dit ter sprake bracht.
'I am completely informed because she mentioned this.'
Indifference context sentence
Ik ben helemaal op de hoogte, maar ik vertel je nu niet alles wat ik weet.
'I am completely informed, but I am not telling you everything I know now.'

Hence, a speaker that provides grounds for her knowledge might have been interpreted as more knowledgeable than a speaker that does not do so.

7.3.3.2 Test 2

Our second questionnaire described in section 6.4.2 tested how coherent our constructed test items were. Participants were given texts like (26) in all six conditions and were asked to judge how compatible the target sentence with the numeral modifier is with the preceding context. They gave their judgements on a 1-7 Likert scale (1: "the sentence does not fit with the preceding context", 7: "the sentence fits well the preceding context").

Similarly to the ignorance Context condition, although the indifference Context items generally received ratings from the upper part of the Likert scale ($mean = 4.389$, $median = 5$), they were judged to be significantly less coherent than the authority Context items ($mean = 5.509$, $median = 6$; $t = -3.138$, $p < .01$), regardless of the NM type ($t = .363$, $p = .721$).

As already pointed out, we did not include any ill-formed, clearly incoherent items in the present questionnaire expected to score low on the Likert scale. So it is likely that participants, trying to make use of the points of the scale they were given in a task that specifically draws their attention to the coherence relation between the target sentence and the preceding text, became more meticulous in their judgements. This might have magnified possible small differences between items, hence, the differences we observed between the two knowledgeability conditions. Interestingly, the feedback we received from participants with respect to the indifference items after the completion of the questionnaire was similar to that for the ignorance condition. A reoccurring remark was that, when the context had an indifferent speaker, the target sentence—irrespective of the numeral modifier—felt like giving specific information and details, which were in conflict with what the relevant speaker had just stated (cf., e.g., *the details are not relevant/interesting*). So this could ex-

plain the attested difference between the two knowledgeability contexts. But, again, why would the target sentence be perceived as giving away more than expected? As with ignorance, we speculate that either the core meaning of the modified numerals, i.e., $[n, \ldots)$, excluding certain values ($< n$) is taken to be too specific/detailed, or the specific indifference implicature associated with these modified numerals, which elaborates that the speaker does not consider n to be relevant and does not consider $[n+1, \ldots)$ to be relevant. According to the composite analysis given in section 7.2, if the latter were the case, there should have also been attested a difference between the two NMs (interaction effect), as the comparative NM is predicted to trigger weaker indifference implicatures. However, the relevant interaction was not found to be significant.

No matter how we interpret the difference under consideration, we can perform a preliminary evaluation of the predictions of the theoretical accounts requiring that the speaker who utters *at least* (in a non-embedding environment) does not have precise knowledge (see relevant predictions in section 7.3.2). Even if there might be independent reasons for the difference between the two knowledgeability conditions, still the indifference Context condition received quite high coherence rates overall, cf. *mean* = 4.389, *median* = 5. One could say that such coherence rates do not represent a severe pragmatic or semantic inconsistency as predicted by Ciardelli et al. (2017), Coppock & Brochhagen (2013b), and Spector (2015) for the superlative items in the indifference Context condition. Consequently, it seems that the results of the present questionnaire as to the indifference Context condition cannot be accommodated by the aforementioned accounts. On the contrary, they are compatible with the composite analysis presented in section 7.2 that captures both speaker indifference and speaker ignorance interpretations of superlative as well as comparative modified numerals (recall also that no difference was observed between ignorance and indifference Context conditions).

To end this section, it is not necessary that the attested difference between the two strong knowledgeability Context conditions will be reflected in the reading task of the eye-tracking experiment, as the present task explicitly instructed participants to zoom in on and judge the coherence relation of the target sentence with the preceding context, which is not the case in the eye-tracking experiment.

7.3.4 Methods

The methods were as described in the corresponding section of chapter 6 (section 6.5). As to the materials regarding the indifference Context condition, we made sure to vary the relevant context sentence across items. Specifically, the indifference context sentences were constructed based on the authority context sentences. As the indifference Context condition was also a knowledgeability condition, the relevant context sentences should include a similar setup of speaker's knowledgeability to that in the authority context sentences. The indifferent speaker would reveal that she is knowledgeable of a certain quantity

and further state that it is not relevant/important in the current purposes of the exchange to elaborate or give the precise number or all the details (see Appendix C for the different ways the indifferent speaker was set up). As already noted, the former part of the indifferent speaker's utterance was what the two knowledgeability conditions had in common.

7.3.5 Results

The preliminary analysis of the data as well as the data clean up have already been described in the corresponding section (6.6) of the previous chapter. Tables 7.1, 7.2, and 7.3 report the raw means of the eye-tracking measures per region per condition, including only values for regions that were fixated after reading later, critical, parts of the individual texts containing the Context and the NM manipulations. For Context they only include the indifference Context condition and the baseline authority Context condition. In the following, I will present the results of the analyses presented in section 6.6 concerning the indifference Context condition, which have not been discussed yet. To remind you, the reference levels of all these analyses were authority for Context and superlative for NM.

7.3.5.1 Region 1, Intro

Although there was no effect of indifference or an interaction effect of NM and indifference on either re-reading measure (i.e., times or probability), there was a marginal positive interaction effect on total reading times (β = .080, SE = .042, t = 1.935, p = .0594): The difference of the indifference Context condition from the baseline authority condition in overall reading times of the Intro tends to be bigger in the comparative items than in the superlative items.

7.3.5.2 Region 2, Speaker knowledgeability context

No significant effect was observed on either measure of interest in Region 2, i.e., probability of regression and probability of re-reading.

7.3.5.3 Region 3, *Wesley*

Region 3, consisting in the first region/word of the target sentence and containing the subject of the sentence, did not exhibit any significant effect in any of the seven measures we looked at.

7.3.5.4 Region 4, *has*

There was no significant effect in Region 4 in any measure according to the analyses we carried out.

Condition		Region		
Context	NM	Region 1 Intro	Region 2 Context	Region 3 *Wesley*
First pass (ms)				
Authority	comparative			279
Authority	superlative			303
Indifference	comparative			311
Indifference	superlative			285
Right bounded (ms)				
Authority	comparative			301
Authority	superlative			307
Indifference	comparative			311
Indifference	superlative			306
Regression path duration (ms)				
Authority	comparative			344
Authority	superlative			334
Indifference	comparative			326
Indifference	superlative			356
Probability of regression				
Authority	comparative		.897	.947
Authority	superlative		.874	.934
Indifference	comparative		.872	.966
Indifference	superlative		.843	.936
Total reading time (ms)				
Authority	comparative	6903		300
Authority	superlative	7072		325
Indifference	comparative	7230		317
Indifference	superlative	6934		330
Re-reading time (ms)				
Authority	comparative	1269		267
Authority	superlative	1235		311
Indifference	comparative	1205		266
Indifference	superlative	1237		283
Probability of re-reading				
Authority	comparative	.365	.266	.386
Authority	superlative	.341	.285	.418
Indifference	comparative	.441	.264	.408
Indifference	superlative	.392	.286	.461

Table 7.1: Means of eye-tracking measures for Regions 1, 2, and 3 in different conditions.

Condition		Region		
Context	NM	Region 4 *has*	Region 5 *that day*	Region 6 NM
First pass (ms)				
Authority	comparative	228	307	235
Authority	superlative	228	327	227
Indifference	comparative	255	296	255
Indifference	superlative	271	298	231
Right bounded (ms)				
Authority	comparative	243	371	246
Authority	superlative	241	385	246
Indifference	comparative	256	392	275
Indifference	superlative	280	393	250
Regression path duration (ms)				
Authority	comparative	315	464	277
Authority	superlative	330	476	320
Indifference	comparative	334	549	385
Indifference	superlative	358	482	318
Probability of regression				
Authority	comparative	.752	.760	.902
Authority	superlative	.700	.746	.860
Indifference	comparative	.695	.652	.857
Indifference	superlative	.695	.691	.867
Total reading time (ms)				
Authority	comparative	282	464	308
Authority	superlative	281	502	329
Indifference	comparative	308	502	351
Indifference	superlative	295	511	321
Re-reading time (ms)				
Authority	comparative	273	402	290
Authority	superlative	259	432	304
Indifference	comparative	275	441	332
Indifference	superlative	250	480	282
Probability of re-reading				
Authority	comparative	.468	.452	.429
Authority	superlative	.503	.488	.402
Indifference	comparative	.578	.540	.483
Indifference	superlative	.520	.512	.411

Table 7.2: Means of eye-tracking measures for Regions 4, 5, and 6 in different conditions.

Condition		Region		
Context	NM	Region 7 *ten people*	Region 8 PP	Region 9 *tattooed*
First pass (ms)				
Authority	comparative	386	367	262
Authority	superlative	399	404	279
Indifference	comparative	397	409	275
Indifference	superlative	380	382	251
Right bounded (ms)				
Authority	comparative	444	390	319
Authority	superlative	438	442	330
Indifference	comparative	457	458	326
Indifference	superlative	433	429	284
Regression path duration (ms)				
Authority	comparative	550	636	903
Authority	superlative	532	631	1116
Indifference	comparative	580	708	1067
Indifference	superlative	582	731	1005
Probability of regression				
Authority	comparative	.794	.818	.479
Authority	superlative	.827	.827	.475
Indifference	comparative	.761	.744	.468
Indifference	superlative	.793	.762	.463
Total reading time (ms)				
Authority	comparative	544	460	343
Authority	superlative	517	528	348
Indifference	comparative	566	545	335
Indifference	superlative	560	508	293
Re-reading time (ms)				
Authority	comparative	421	314	360
Authority	superlative	432	394	313
Indifference	comparative	404	426	289
Indifference	superlative	460	387	255
Probability of re-reading				
Authority	comparative	.429	.424	.270
Authority	superlative	.319	.426	.254
Indifference	comparative	.489	.425	.285
Indifference	superlative	.440	.448	.256

Table 7.3: Means of eye-tracking measures for Regions 7, 8 and 9 in different conditions.

7.3.5.5 Region 5, *that day*

The analyses showed no signifiant effect on any measure in Region 5 either.

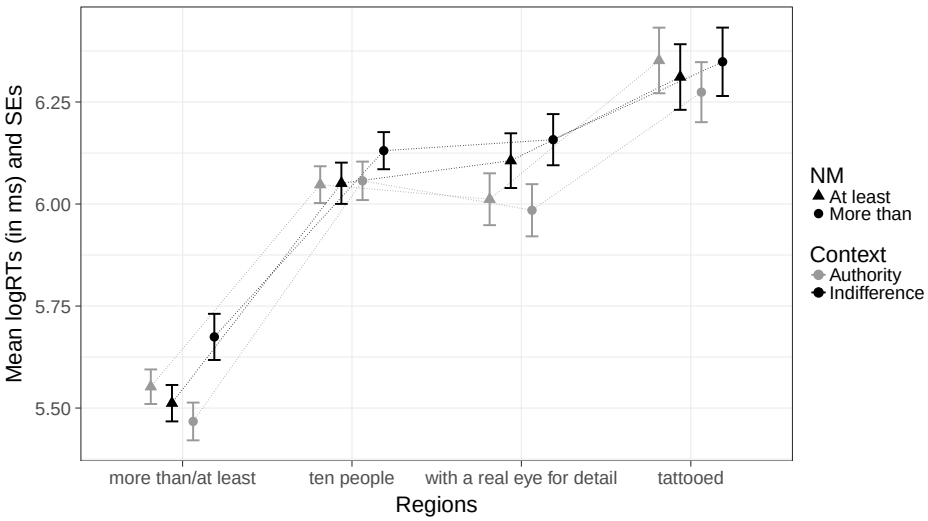

Figure 7.1: Regression path reading times for each region from the numeral modifier region onward per condition.

7.3.5.6 Region 6, NM

Of all measures we analyzed for Region 6, we found a significant positive interaction of NM and indifference Context in regression path duration such that participants would slow down their reading of the numeral modifier and regress back to previous regions in the indifference Context condition as compared to the baseline authority condition significantly more when the numeral modifier was *meer dan* ('more than') than when it was *minstens* ('at least') (β = .251, SE = .088, t = 2.843, p < .01). In other words, the effect of Context (authority vs. indifference) was different for comparative items than for superlative items. A post-hoc analysis for the separate numeral modifier conditions revealed that there was an effect of indifference Context only in the comparative condition: Indifference Context items were read slower including regressions to previously read regions than authority items in the comparative NM condition (β = .222, SE = .080, t = 2.794, p < .05), you can clearly see this difference in figure 7.1. That is, it seems that the observed interaction effect is due to the the effect of longer regression path reading times for indifference

items in the comparative condition. This analysis also showed that indifference Context items were significantly different from ignorance Context items, too, in the comparative condition (β = .211, SE = .081, t = 2.595, p < .05). These findings come to confirm the pattern visible in figure 7.1, representing how higher the relevant reading times are in the indifference condition of comparative items than anything else.

The same interaction effect showed up in total reading times (see figure 7.3), however it approached but did not reach significance (β = .165, SE = .092, t = 1.802, p = .080).

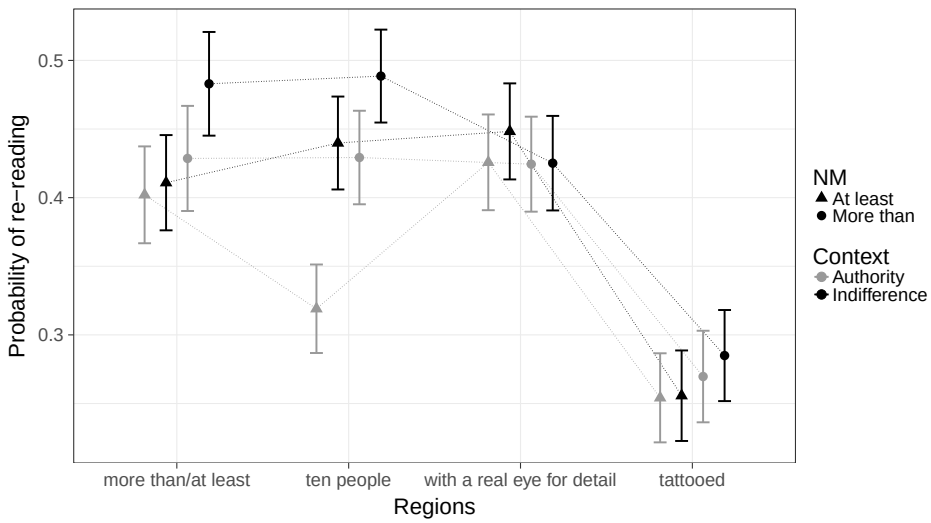

Figure 7.2: Probability of re-reading each region from numeral modifier onward per condition.

7.3.5.7 Region 7, *ten people*

In the region of *ten people* that comes to complete the modified numeral phrase and its interpretation, where we found significant effects of ignorance Context in both eye-tracking experiments (chapters 5 and 6) as well as a significant interaction such that the effect of ignorance is bigger for superlative than for comparative items (see chapter 6), we find a similar picture as far as the indifference Context condition is concerned. Specifically, the aforementioned main and interaction effects were manifested in the measure of the probability of re-reading the region in question and no other significant effect arose in any other measure. As to indifference, the only significant effect that we observed appeared in that same measure and was a positive effect of indifference Context

(β = .600, SE = .216, t = 2.779, p < .01): Readers were more likely to re-read the region of *ten people* of indifference items than that of authority items in the baseline superlative NM condition (see relevant difference in figure 7.2). Contrary to the ignorance-related results though, although the interaction between the NM and the indifference Context condition had a similar direction, it was not found to be significant (β = −.341, SE = .297, t = −1.147, p = .252).

7.3.5.8 Region 8, PP

There was no significant main or interaction effect in any measure, though a positive interaction effect approaching borderline significance (β = .146, SE = .078, t = 1.864, p = .066) was observed in total reading times (see figure 7.3).

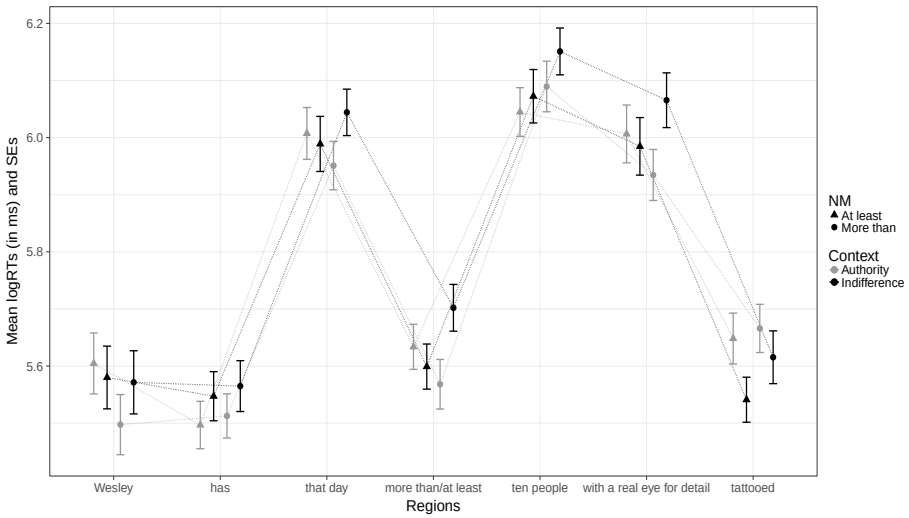

Figure 7.3: Total reading times for each region of the target sentence per condition.

7.3.5.9 Region 9, *tattooed*

No significant effect of indifference Context or relevant interaction was found in the last region of the target sentence.

7.3.6 Discussion

Effects in Region 7. The indifference effect found in this region is similar to the effect found for ignorance in this very region (see sections 6.6.7 and 6.7).

That is, there was a positive effect of indifference Context in the probability of re-reading measure, indicating that comprehenders are more likely to re-read the numeral phrase (*ten people*, Region 7) modified by the superlative modifier in an indifference rather than in an authority context. Given that we have concluded that the corresponding, almost identical, ignorance Context effect is to be attributed to the (indirect or on-line) derivation of the specific speaker ignorance implicature associated with the superlative modifier in the target sentence, it seems that a similar explanation could hold for the observed indifference Context effect. Namely, readers revisit more often Region 7, where the interpretation of the whole modified numeral phrase is completed, either (i) in order to derive the specific indifference implicature, which is not entailed by the context, or (ii) because, after having already derived that implicature (rapidly) in their first pass, they figure that the resulting reading is too specific compared to the underspecified speaker indifference entailed by the context (*there is no certain number that the speaker considers relevant*). Remember that we did obtain lower compatibility scores in the second questionnaire (section 7.3.3.2) for the indifference compared to the baseline authority Context condition, and also participants remarked that when the context had an indifferent speaker the target sentence felt too specific and detailed than expected given the particular context, irrespective of the numeral modifier. This is in line with the explanation in (ii) and analogous to what have been found and discussed for ignorance in the previous chapter. Therefore, the finding in question manifesting the availability of a specific indifference implicature with the superlative modifier seems to be in support of a unified account of speaker ignorance and speaker indifference involving the same implicature mechanism, such as the account outlined in section 7.2.

Furthermore, the attested difference between indifference and authority contexts in the superlative NM condition further evaluates, and more precisely invalidates, the predictions stemming from a number of existing accounts obligatorily deriving speaker ignorance with unembedded uses of *at least*. As already said, if anything, we would surely expect a similar pattern between the two (strong) knowledgeability conditions (i.e., indifference and authority, see section 7.3.3.1) according to these accounts. Specifically, the inquisitive semantics accounts by Ciardelli et al. (2017) and Coppock & Brochhagen (2013b) predict a similar incompatibility to arise between the target sentence with the superlative modifier and the context sentence of the two knowledgeability conditions. The same incompatibility is predicted given Spychalska's (2015) account in the two knowledgeability conditions due to the violation of the assertibility condition of *at least* requiring that the speaker uttering *at least* should entertain two possibilities in her information state, which is not the case for the speaker in the knowledgeability contexts of authority and indifference. Lastly, the different pattern observed between the two knowledgeability contexts in the superlative condition goes against Spector's proposal in Spector (2014, 2015), which predicts a contradiction in the two Context conditions between the speaker's information state as set up by the relevant context sentences and that obliga-

torily signaled by the use of *at least*.

No significant interaction effect was observed between the NM and the Context factors as to indifference, contrary to ignorance. Consequently, we can draw no conclusions as regards the strength difference in specific indifference implications predicted by the composite account for comparative and superlative modifiers. This is actually in line with the preliminary data we obtained from the questionnaire reported at the end of section 7.2, where no significant difference between the two NMs was observed either as to indifference implicature rates.

Effects in Region 6. There seems to be quite a severe disruption of the comprehension of comparative items in the indifference Context condition. This disruptive effect is found to be significant in Region 6, the region of NM, in regression path duration, and it becomes marginal in total reading times of Region 1 (Intro) and then also of Region 6, and likewise in the later, spillover Region 8 (*with a real eye for detail*), see figure 7.3. Although the relevant interaction effect observed has the same direction of the interaction effect found to be significant in re-reading times of Region 8 for ignorance, it seems to be of a different source and nature as it shows up very early in the eye-movement record. The significant interaction effect attested in regression path duration of Region 6 indicates that once participants encounter the comparative modifier *meer dan* in the indifference condition they slow down and regress to previously read material. Why do they do so?

As said, the location and timing of the effect tell us that this cannot be due to the same reason as the relevant effect found for ignorance, because the latter occurred much later in the eye-movement record (see discussion of effects in Region 8 in section 6.7). Hence, it is not to be explained as an indication of the derivation of a weaker specific speaker indifference implicature with the comparative modifier as compared to the superlative modifier and/or of a speaker indifference implicature that is different in nature. Nor should it be explained in terms of a Manner implicature such that *minstens* is a better cue to speaker indifference than *meer dan* is, and once readers encounter the latter modifier they wonder why the speaker used that rather than the former modifier in the particular context. Such an interpretation also leaves unexplained why the relevant effect occurs in the regression path duration measure and, thus, what makes readers regress back to previous regions in the comparative NM condition. More importantly, this interpretation goes against our intuition that sometimes the comparative modifier can be better/more appropriate than the superlative modifier at expressing speaker indifference as to a precise degree or amount, cf. the relevant contrasts in (30) as well as in (31) below, where the utterances with the comparative modified numeral feel more felicitous than the utterances with the superlative modified numeral (see further discussion in section 7.4):

(30) **A:** Did the water freeze?
B: No, it was more than 0 degrees Celsius.
vs. No, it was at least 0.1 degree Celsius.

(31) **A:** Are you a millionaire?
B: Yes, my wealth is more than 1 million dollars.
vs. Yes, my wealth is at least 1 million dollars.

So, still, what makes participants slow down once they read *meer dan* in the indifference condition and also regress to previous material, possibly at the Intro, as suggested by the relevant marginal interaction effect in this region? Here's an idea hinted at by feedback I obtained on the items: *Meer dan* seems to have quite a strong evaluative effect in the indifference condition that *minstens* does not have.[9] The speaker in the indifference condition states that the details, that is, mentioning a precise number n, are trivial, and all she regards as relevant is that n exceeds (or reaches for the superlative NM) a minimum value/threshold m_1, which is higher than a previously mentioned threshold m_0 appearing to be the standard. To illustrate with the example item from section 7.3.1:

(32) **Translated example item**

Intro
Wesley runs his own tattoo parlor, which he enjoys a lot. It's usually very busy and he tries to tattoo eight customers per day. Last Thursday, he was very busy.

Speaker indifference context
'I know exactly how busy it got, but it's not so interesting to elaborate on that.'

Target sentence
That day, Wesley tattooed at least/more than ten people with a real eye for detail.

Let n be the number under discussion known by the indifferent speaker. m_1 = 10, i.e., the number modified in the target sentence, while m_0=8, the value mentioned as a standard in the Intro. Thus, the indifferent speaker states that n is not relevant, while all she regards relevant is that n exceeds/reaches the value 10, which is higher than the standard 8. The use of the comparative modifier *meer dan additionally* signals that n being higher than 10 (m_1) is noteworthily *a lot* given the standard, i.e., 8 (m_0). The superlative items in the same condition lack this evaluative effect, as they merely state the relevance of reaching threshold 10 (m_1), which is higher than the standard 8 (m_0). Given that, it seems that we can interpret the effect we observe in Region 6 as follows: As soon as readers see *meer dan* in the indifference condition and also the number phrase coming up to the right, they regress possibly to the Intro including the

[9] Many thanks to Maartje Schulpen for discussion on that.

standard m_0 to see and check whether n being higher than m_1 was a lot higher than the standard. Note that the effect in question suggests that the evaluative effect of *meer dan* is not/less present in the baseline authority condition. The authority condition does not include the cue that the speaker is indifferent to the precise number, to the contrary it is implied that the precise number n is relevant. Hence, it seems that the availability of the *indifference* information gives a boost—or even is related—to the evaluative effect of the comparative modifier.

Another plausible explanation, which accords with our intuition as to examples (30) and (31), would be that the early processing penalty in the eye movement record at the expense of the comparative condition is due to the online derivation of an indifference implicature that is more robustly associated with the comparative modifier (see more discussion on this in section 7.4.1). This, too, is incompatible with the predictions of the extended composite theory.

7.4 Recap and discussion

In this chapter, we considered cases where unembedded uses of modified numerals might signal that the speaker is indifferent to a precise cardinality or that all that matters or is relevant at the current context is that the minimum value compatible with the modified numeral is true. We have indications that this signal is in fact yet another inference that an utterance with a modified numeral can trigger (besides partial variation and speaker ignorance, that is), which is derived via some pragmatic reasoning. In order to account for this inference and in line with literature on speaker indifference effects in other empirical domains, we suggested to extend the composite theory proposed (in chapter 6) based on findings from the previous chapters studying variation and ignorance effects. The resulting account derives speaker indifference inferences, too, as primary Quantity implicatures, and predicts a strength difference between superlative and comparative modified numerals, with the former being associated with stronger speaker indifference effects. However, both the data we collected via a preliminary off-line inferential task and the relevant results of the eye-tracking experiment 2 do not confirm such a difference in implicature strength between the two types of numeral modifiers, though the latter experiment does bring to light other differences between them (possibly having to do with the particular setup of our texts). Although the indifference items do pattern with ignorance items in the superlative condition in experiment 2 in support of a unified account of the respective inferences, no difference is attested between superlative and comparative items in the direction predicted by the composite theory. This is surprising and suggests that the composite theory is not extendable to and does not work for speaker indifference inferences. That is, it seems that speaker indifference inferences of numeral modifiers do not pattern with the corresponding speaker ignorance inferences and should be

accounted for differently. The finding in question will be further considered in the next sections.

Specifically, in the following, I discuss certain issues that arise with regard to speaker indifference and the relevant contexts that were constructed in eye-tracking experiment 2 in order to test for such inferences.

7.4.1 Loose ends of extended composite theory

That the extended composite theory seems not to be tenable is also suggested by the fact that sometimes comparative modified numerals feel to be better at signaling speaker indifference than their superlative counterparts, as already noted in the previous section. A possible cause for that might be that speaker ignorance is interfering with—or prevailing in—the interpretation of an *at least* utterance, even in (31) where the speaker's own wealth is under discussion and even though in both (30) and (31) the utterance with the modified numeral constitutes an answer to a polar question (remember that polar questions have been found by Westera & Brasoveanu, 2014 to weaken speaker ignorance interpretations of modified numerals in the relevant responses). This interference of speaker ignorance is perhaps evident in the student's utterance in (33), where the underlined *at least* sentence appears to be odd.[10]

(33) **Teacher:** As I said, to qualify for the next round of the geometry competition, you must draw a polygon that has at least four sides.
Student: I qualify! Look! I drew a rectangle; ?<u>it has at least four sides</u>.

To elaborate, the *at least* sentence in the student's utterance is supposed to be acceptable on an indifference reading, i.e., all the student considers to be relevant is that her figure has 3 or more sides and it is not relevant to specify the exact number of sides. The fact that this sentence might feel degraded suggests that it is interpreted as signaling speaker ignorance. However, Mendia (2016b) provides a similar example as fully acceptable, see (34), intending to illustrate that speaker ignorance inferences of superlative modifiers are cancellable, cf. (*i*)*n fact*-sentence.

(34) **Context:** Bill has four kids. Yesterday he saw a sign at a supermarket: "Huge sales and discounts for parents. To qualify, you must have at least three kids." After reading it, Bill reasoned as follows: "I qualify, <u>I have at least three kids. In fact, I have four.</u>" (Mendia, 2016b, p. 18)

Although the overall picture is not very clear, it seems to suggest that while superlative modifiers signal stronger speaker ignorance, comparative modifiers are better at expressing speaker indifference. It might be the case that, although both types of numeral modifiers trigger speaker ignorance *and* speaker indifference, speaker ignorance with superlative modifiers is in the process of becoming conventionalized (see Grice, 1975, p. 58). This can explain the murkiness

[10]Thanks to Maria Barouni for bringing up this point.

of the judgements concerning (33) and (34), but also why comparative modifiers appear to be preferred over superlative modifiers for expressing speaker indifference (cf. (30) and (31)) or to have a stronger speaker indifference implicature. Along these lines, as already mentioned, we can maintain that the interaction effect first appearing in Region 6 in early measures and the opposite one exhibited in Region 7 in later measures constitute a manifestation of the strength difference in indifference effects between the two numeral modifiers: Comparative modifiers trigger more robust indifference inferences compared to superlative modifiers, hence the timing difference. Crucially, I should make clear that it is not necessary to reject the extended composite theory in light of our overall findings in eye-tracking experiment 2 as long as we take the ignorance conventionalization scenario into account. That is, the extended composite theory is still tenable if we take it as the onset of the aforementioned conventionalization process.

7.4.2 Alternative function of indifference condition in eye-tracking experiment 2

There is a possibility that the particular indifference condition we have set up in order to test for indifference inferences actually tests for a different type of inference modified numerals give rise to in non-embedding environments, the so-called granularity-based scalar implicatures researched by Cummins et al. (2012).

In particular, Cummins et al. (2012) have shown experimentally that, contrary to what has been argued in the relevant literature (Fox & Hackl, 2006; Krifka, 1999), unembedded uses of *at least* and *more than* by a knowledgeable speaker can generate upper-bound interpretations (via scalar implicatures), if granularity is taken into account. Specifically, *at least n* and *more than n*, when used by a cooperative and knowledgeable speaker, can trigger the implicature that *at least m* and *more than m*, respectively, are false, where m is a stronger alternative to n and *the coarsest granularity level expressed by m is at least as coarse as that expressed by n* (Cummins et al., 2012, p. 143). To illustrate, although (35) does not give rise to the strengthened interpretation in (35-a), it does trigger the implicature in (35-b) or (35-c), because the relevant m is the next point on the assumed scales < *more than* $50,000$, *more than* $60,000$ > (numerals as scale points on the scale of granularity 10,000) and < *more than* $50,000$, *more than* $55,000$ > (numerals as scale points on the scale of granularity 50,000), respectively.

(35) Kalamata has at least/more than 50,000 inhabitants.

a. $\not\rightsquigarrow$ *Kalamata does not have at least/more than 50,001 inhabitants* + (35) → *Kalamata has exactly 50,000/50,001 inhabitants*

b. $\rightsquigarrow$ *Kalamata does not have at least/more than 60.000 inhabitants*

c. $\rightsquigarrow$ *Kalamata does not have at least/more than 55.000 inhabitants*

In this light, instead of assuming that the whole set of (stronger) alternatives associated with each type of numeral modifier (see (1)) is used in a particular way by the indifferent, but not by the authoritative, speaker in our experiment, I am going to consider the following possibility: (Only) values that belong to a scale of a certain granularity are activated in the indifference Context condition and then used in a standard scalar implicature-based reasoning, like that described by Cummins et al. (2012) for modified numerals. Let us see how this story would go.

Given the mention of a certain value (standard, m_0) at the Intro of the texts, the indifferent speaker, who is knowledgeable of the number under discussion—standing in some contrast to or comparison with the standard value (m_0), points out (in the target sentence) another value, which is worthy of mentioning in order to refer to the number under discussion. Immediately, via the mention of the second value (m_1), which crucially is always a round and imprecise number (see section 6.2.1), the speaker establishes the level of granularity of the scale to be associated with the numeral in the target sentence. The value mentioned in the Intro (standard, m_0) can be taken to sit on that same scale, thereby giving some extra cues as to the type of scale to be considered.

In this context, the choice of the round number modified in the target sentence by a knowledgeable and cooperative speaker in the indifference Context condition triggers a number of implicatures as to the upper bound, depending on the granularity level at which the numeral in question is interpreted. For instance, the indifferent speaker in our familiar example item, given in (36) (in English), might be taken to convey that it is false that Wesley tattooed at least/more than 15 people, assuming a scale of granularity 5, or it is false that he tattooed at least/more than 20 people, assuming a scale of granularity 5 or 10, etc.

(36) Wesley runs his own tattoo parlor, which he enjoys a lot. It's usually very busy and he tries to tattoo eight people per day. Last Thursday, he was very busy. I know exactly how busy it got, but it's not so interesting to elaborate on that now. That day, Wesley tattooed at least/more than ten people with a real eye for detail.

Or, in order to also demonstrate a case where the number mentioned in the Intro might help better determine the granularity of the scale to be taken into account in the target sentence, in (37) (see Appendix C for original item), the (indifferent) speaker most possibly is taken to suggest that it is false that An copied at least/more than five hundred pages on Monday morning. That is, the numeral in the target sentence is interpreted at the granularity level 100.

(37) An is a secretary at a bank and has to copy many documents every day. Most of the days, she copies three hundred pages in total. Monday afternoon, there would be an important meeting. She informed me about how much she had to do, but I will not tell you the details now. On Monday morning, An copied at least/more than four hundred pa-

ges for her colleagues.

As has become clear thus far, it seems that there is another way of looking at the indifference Context condition of experiment 2. Namely, we could take the relevant speaker we have set up to convey scalar implicatures like the above rather than weaker Quantity implicatures such that the speaker does not consider it is relevant to mention n and does not consider it is relevant to mention $[n+1,\ldots)$ (where n the minimum value compatible with the modified numeral in the target sentence). On such an interpretation, we could take the observed processing penalties associated with the indifference Context condition to reflect the on-line calculation of the granularity-based scalar implicatures.

But what would prevent such implicatures to arise in the authority Context condition? In that case too, the speaker is knowledgeable and cooperative, and mentions a round number that is relevant and can give rise to the same implications. The only difference between the two knowledgeability contexts is that the speaker in the indifference condition further specifies that it is not interesting/relevant to give all the information/details she knows or that she will just give an approximation, although she has full knowledge. This very specification could be assumed to constitute a bias toward drawing the scalar implicatures, as it sets up an imprecise/coarse QUD prompting for the activation of a coarse-grained scale to be associated with the numeral in the target sentence. As a result, more granularity-based scalar implicatures are calculated in the indifference rather than in the authority knowledgeability condition, which lacks such a prompt, hence the attested differences.

Now, how could we accommodate in the above story the (robust) processing penalty we find to be associated with the comparative NM condition (positive interaction effect)? The relevant effect hindering comprehension as early as the numeral modifier is encountered (for the first time) might indicate that more (robust) granularity-based implicatures are triggered by the comparative modifier as compared to its superlative counterpart. In other words, the scalar implicature in the comparative condition is computed when readers see for the first time the modifier and the round number ensuing, and regress possibly to the number in the Intro (standard value, m_0) to help themselves construct the right granularity scale, the basis of the scalar implicature. After that readers proceed with reading, which seems to be affected in this condition up to the penultimate region with the PP. In the superlative condition, they are likely to generate a scalar implicature quite later in the eye-movement record and the target snetence, that is, at the region with the number (Region 7: *ten people*), as evidenced by the effect in the probability of re-reading that region in the indifference condition as compared to the baseline authority condition. Why this implicature would be more robustly available with comparative than with superlative modifiers will not be addressed in the present thesis.[11]

[11] A possible direction for an answer could be suggested by Cummins et al.'s (2012) finding that (unprimed) *more than n* utterances implicate a higher upper bound than the corresponding *at least n* utterances do.

To end, this chapter set out to probe speaker indifference effects of numeral modifiers from both a theoretical and an experimental perspective. Specifically, it investigated the following questions from the Introduction (chapter 1):

- ▷ Is speaker indifference an available implication of modified numerals? How could we capture speaker indifference and how does it relate to speaker ignorance? Does it manifest itself in incremental interpretation of modified numerals, also in comparison with speaker ignorance effects?

By means of a number of linguistic diagnostics and an off-line study, I argued that numeral modifiers like *at least* and *more than* give rise to speaker indifference effects. In line with studies on speaker indifference in other empirical domains, I proposed to account for speaker indifference in a similar way to speaker ignorance, by extending the composite theory put forth at the end of chapter 6 based on the obtained ignorance-related results of eye-tracking experiment 2. The predictions of this theory were tested by an extra Context condition (speaker indifference condition) included in eye-tracking experiment 2. We did not attest the same patterns for the indifference condition as for the ignorance condition, which was considered a challenge for the applicability of the extended composite theory. This was anticipated given the preliminary off-line data we obtained revealing no strength difference between superlative and comparative modifiers as to indifference interpretations. In light of strong intuitions revealing that comparative modifiers are better cues to speaker indifference than superlative modifiers are, I suggested that we can still maintain the extended composite theory as long as we take ignorance conventionalization into account for superlative modifiers. That is, although both types of numeral modifiers trigger speaker ignorance and speaker indifference, speaker ignorance with superlative modifiers is in the process of becoming conventionalized and as a result comparative modifiers have been led to take over and sound better at signaling speaker indifference in non-embedding environments. No matter how we interpret the findings regarding the indifference Context condition of eye-tracking experiment 2 (see previous section), this chapter adds to the overall picture of the pragmatics of numeral modifiers, setting the stage for further research on certain issues (such as ignorance conventionalization, and the relation of what we have called speaker indifference to the granularity-based scalar implicatures of modified numerals).

The following chapter concludes the present thesis, reflecting on the findings of all previous chapters and on how these contribute to a better understanding of the pragmatics of numeral modifiers as well as to the experimental pragmatics research in general.

CHAPTER 8

Conclusion

8.1 The pragmatics of numeral modifiers

This thesis is a study on the pragmatics of numeral modifiers. Nouwen (2010) divides the various truth-conditionally equivalent numeral modifiers into two distinct classes: Class A modifiers are modifiers that lack ignorance effects, while class B modifiers display these effects. The dominant take in the relevant literature accounts for ignorance effects in terms of pragmatic mechanisms. Only very recently has this distinction been questioned by a small number of studies. Ciardelli et al. (2017) and Westera & Brasoveanu (2014) (see also unpublished work by Mayr & Meyer, 2014) have argued that class A modifiers too can trigger ignorance effects when they are used as an answer to a precise question under discussion. Of these two studies only the former accounts for these effects with both class B and class A modifiers. Moreover, numeral modifiers can also signal speaker indifference as well as variation effects. While the latter effects have scarcely been investigated, the former have remained completely unexplored in the domain of numeral modifiers.

In this context, this dissertation probes experimentally all three types of inference, viz., variation, speaker ignorance, speaker indifference, with both class A and class B numeral modifiers aiming to find out where one should draw the dividing line between the core meaning and the pragmatic meaning of each class of modifiers, and how different the two classes are, with a main focus on their pragmatics. Doing so will help provide a more adequate answer to the question whether all different numeral modifiers language provides convey actually the same meaning.

In the remainder of this chapter, I review the findings of the thesis in relation to the research questions I posited in the Introduction chapter and I further discuss the theoretical and psycholinguistic benefits of this work and what still needs to be done.

8.2 Chapter summary

Chapter 1 introduced the individual research questions that guided this thesis towards achieving its overarching goal, as described in the previous section:

- ▷ What is the likelihood of drawing a variation inference? Do different (truth-conditionally equivalent) modified numerals give rise to variation effects to the same extent? What is the semantic/pragmatic status and strength of variation effects? What type of variation do modified numerals involve and what are the implications for their underlying mechanism of derivation?
- ▷ How are ignorance inferences accessed in incremental interpretation of modified numerals? What are the insights into the nature and strength of speaker ignorance inferences and into the underlying mechanism of derivation? How is this mechanism different across different modified numerals? How does the mechanism responsible for speaker ignorance effects relate to that responsible for variation effects?
- ▷ Is speaker indifference an available implication of modified numerals? How could we capture speaker indifference and how does it relate to speaker ignorance? Does it manifest itself in incremental interpretation of modified numerals, also in comparison with speaker ignorance effects?

Chapter 2 reviewed the extensive theoretical literature on ignorance effects of modified numerals as well as the few existing accounts of their variation effects, with a focus on the distinction between superlative and comparative modifiers, which are class B and class A modifiers, respectively. The aim of this chapter was to highlight the aspects of these accounts that would be crucial for the experimental work I conducted and reported on in the chapters that followed.

Chapter 3 tackled the first set of questions (see above), concerning variation effects of numeral modifiers. It presented novel experimental data obtained by means of a series of off-line experiments conducted in Dutch that tested both class B (*minstens* 'at least', *n of meer* '*n* or more', *maximaal* 'maximally', *n of minder* '*n* or less/fewer') and class A (*meer dan* 'more than', *minder dan* 'less/fewer than') numeral modifiers. I tested the likelihood and strength of variation effects of numeral modifiers in the scope of a universal nominal quantifier, and also set out to find out which type of variation effects numeral modifiers trigger, when in the scope of universal nominal or modal quantifiers:

Underspecified or specific partial variation. Taken together, our experiments show that superlative, disjunctive, as well as comparative numeral modifiers trigger non-obligatory pragmatic variation implications. It was further shown that both superlative and comparative modifiers trigger specific partial variation effects, i.e., the variation interpretation includes an inference about the minimum value n compatible with the modified numeral. A difference in acceptability ratings noticed between comparative and superlative modifiers was taken to suggest a difference in the strength of the specific variation effects the two types of numeral modifiers trigger, with comparatives being associated with weaker specific variation effects than superlatives. Crucially, the finding as to the type of variation triggered by numeral modifiers yielded implications for the underlying mechanism of derivation. In particular, it indicates that $[n]$ and $[n+1,\ldots)$, where n is the minimum value compatible with the modified numeral, are the alternatives fed into the relevant pragmatic mechanism.

Chapter 4 detailed the extant experimental studies on speaker ignorance inferences of superlative and comparative modifiers, setting the stage for the experimental investigation to follow in chapters 5 and 6. As discussed, the findings of the majority of the studies provide evidence that ignorance inferences are available only with superlative modifiers and have a pragmatic rather than a semantic status. Special attention was given to Westera & Brasoveanu's (2014) study, the first one to argue and show that comparative modifiers too can trigger ignorance effects in certain environments. It was further pointed out that the existing evidence under consideration is mostly indirect, while the very little available direct evidence was argued to be rather controversial. Based on this, it was concluded that a more detailed direct study of ignorance inferences was needed, which would look into the incremental interpretation of numeral modifiers. Such an investigation could shed interesting light on the underlying processes responsible for the derivation of speaker ignorance inferences with both superlative and comparative modifiers, and would particularly help adjudicate among the various pragmatic proposals of ignorance.

Chapters 5 and 6 investigated speaker ignorance inferences during incremental interpretation, specifically addressing the second set of research questions of those listed above. Chapter 5 reported on an eye-tracking experiment, which employed the reading paradigm and was conducted in Dutch. This experiment tested unembedded and embedded occurrences of the superlative modifier *minstens* ('at least') in a context with a partially knowledgeable (ignorant) speaker vs. a fully knowledgeable (authoritative) speaker. The context setup was basically inspired by Breheny et al. (2006), which—among a large number of subsequent studies—demonstrated the context-sensitivity of scalar implicature computation, against a default approach to the computation of implicatures. In light of this robust finding, the idea behind our Context manipulation was the following: The contexts with the ignorant speaker would favor the calculation of a specific ignorance implicature in the target sentence with *minstens* ('at least'), while such implicatures would be less likely to arise in the contexts with the authoritative speaker. Speaker ignorance interpretations of embed-

ded uses of the superlative modifier arise according to the the majority of the theoretical accounts via wide scope movement of the superlative modifier in relation to the present operator. So we tested the superlative modifier also in embedding modal contexts in order to find out whether traces of wide scope movement of *minstens* could be detected by the eye-tracking means. As the various accounts of speaker ignorance effects are mainly theoretical accounts with no psycholinguistic claims, we maintained a neutral approach to the processing of the superlative modifier and the on-line accessing of its specific speaker ignorance inference, by considering different processing possibilities. We found a processing penalty in the re-reading probability of the region with the numeral phrase (e.g., *six people*) following the numeral modifier region in ignorance contexts. No interaction of Context and Verb factors was attested. That is, we found no evidence indicating that ignorance with *at least* in interaction with a universal modal involves an extra operation, like covert movement.

Our overall findings were taken to suggest that speaker ignorance effects of superlative modifiers are non-obligatory, context-dependent implications of the specific ignorance type and are derived via a pragmatic mechanism like that specified by the Quantity-based Gricean accounts of ignorance. I concluded that the derivation of speaker ignorance interpretations is associated with a processing cost, displayed in late measures. By means of this experiment, I answered the first part of the set of questions concerning ignorance inferences (see above). However, I further considered a number of possible alternative explanations of the attested effect attributed to speaker ignorance derivation, which necessitated carrying out a follow-up experiment.

Chapter 6 presented this follow-up experiment, which resolved certain challenging issues and aimed to find out whether the effect observed in the previous eye-tracking experiment was to be associated with the derivation of a specific ignorance implicature. More specifically, based on native speakers' feedback, I considered the possibility that the information of the target sentence with *minstens* ('at least') was too specific, and hence infelicitous, given the preceding ignorance-statement by the ignorant speaker, and this caused the processing penalty we attested. I argued that what could be taken to be specific about the target sentence with the superlative modifier was (i) the core meaning of *minstens* n ('at least n'), excluding the values lower than n, or (ii) the specific ignorance implicature associated with *minstens*. In order to find out whether the observed Context effect was to be attributed to the derivation of the specific ignorance implicature—either on-line or off-line through the disruption of the infelicity due to (ii)—or was the result of the disrupting infelicity due to (i), I also tested the comparative modifier *meer dan* ('more than'). The comparative condition would help tease apart these possibilities, as, on the one hand, *meer dan* has a core meaning as specific as that of *minstens*, and, on the other hand, it was assumed to be associated with a less robust specific ignorance implicature than *minstens* is. Hence, it was hypothesized that if the Context effect observed in the previous experiment was due to the infelicity caused by the specific core meaning of the modified numeral given the preceding ignorance

context, the two numeral modifier conditions should behave the same. However, the two numeral modifiers should exhibit a different processing profile, if the observed Context effect was due to the (on-line or off-line) derivation of a specific ignorance implicature.

We replicated the Context effect of the first eye-tracking experiment: Readers tended to re-read the region of the numeral phrase (*ten people*) more often when they were in an ignorance as opposed to an authority context. Crucially, they were more likely to do so in the superlative than in the comparative condition. In general, the processing profile of the comparative condition was different from that of the superlative condition. A processing penalty at the expense of the comparative condition of ignorance items was further observed at re-reading times of the region after the modified numeral phrase (e.g., *with a real eye for detail*). Given these findings, I concluded that the Context effect attested in the first eye-tracking experiment and replicated by the second one should be attributed to the derivation of a non-obligatory specific speaker ignorance implicature at the numeral phrase modified by *minstens* in a speaker ignorance context—manifested either on-line or indirectly through the infelicity of the resulting interpretation of the target sentence given the preceding ignorance utterance.

The processing penalty observed at the prepositional phrase after the modified numeral was taken to be an indication that another pragmatic process was taking place at a later point of the target sentence at the expense of comparative items (in ignorance condition). Either a weaker specific ignorance implicature and/or of a different nature was generated with the comparative modifier as compared to the superlative modifier (manifested on-line or indirectly, for the same reason as for the superlative modifier), or readers, on second thought, did the following Manner-like reasoning after the encounter of the comparative modifier in an ignorance context: Why did the speaker/author not use *minstens* ('at least') instead, which would have been a better cue to ignorance? This results in the derivation of a Manner implicature, attested on-line. It was finally claimed that, given the Manner-based explanation, *minstens* is a better cue to ignorance because it is associated with a stronger ignorance implicature.

Furthermore, I proposed to capture the attested strength difference in specific ignorance inferences between the two types of numeral modifiers as well as their non-obligatory and context-dependent status by the so-called composite theory. This theory basically incorporates Nouwen's (2015) account for superlative modifiers and a standard Quantity-based account for comparative modifiers. That is, superlative and comparative modifiers are associated with the same set of alternatives, illustrated below:

(1)

$$\left\langle \begin{array}{lllll} \ldots & [n+2] & [n+1] & [n] & \\ & & & & [\text{n}, \ldots) \\ \ldots & [\text{n}{+}3, \ldots) & [\text{n}{+}2, \ldots) & [\text{n}{+}1, \ldots) & \end{array} \right\rangle$$

(n is the minimum value compatible with the modified numeral, $[n]$ stands for the proposition with *only/exactly* n and $[n, ...)$ stands for the proposition with the modified numeral). Importantly, although the two types of numeral modifiers share the same alternatives, those are partly different as to their source. The *exactly* alternatives are presupposition-based for superlative modifiers and are fed into a Manner-based mechanism, while for comparative modifiers those alternatives come about via a Horn scale and are used in a standard Quantity-based reasoning. The rest of the alternatives come about via the Horn number scale and are subjected to a Quantity-based process for both superlative and comparative modifiers. Depending on whether numeral modifiers appear in embedding environments (e.g., universally quantified ones) or non-embedding environments, the relevant pragmatic mechanisms output specific partial variation and specific speaker ignorance implicatures, respectively. With this proposal we concluded our investigation of speaker ignorance effects of modified numerals, providing also an adequate answer to the second set of research questions concerning ignorance (see above).

To end, chapter 7 aimed to tackle the last set of questions, relating to speaker indifference. By means of a number of linguistic diagnostics and an offline study, I argued that superlative and comparative numeral modifiers can also give rise to speaker indifference effects such that the speaker considers that it is not relevant to specify the exact number under consideration. In line with studies on speaker indifference from other empirical domains, I proposed to account for speaker indifference in a similar way to speaker ignorance, by extending the composite theory to uniformly capture both types of inferences as (primary) Quantity implications. In fact, I kept the mechanism for each type of modifier as described above, but replaced the belief operator, $\Box_{Bel}$, responsible for the derivation of ignorance implications, with a more general operator, i.e., $\Box_{Assert}$. $\Box_{Assert}\ p$ stands for *the speaker believes p to be true and considers p to be relevant*. Hence, given an utterance $\Box_{Assert}[n, \ldots)$, the resulting ignorance and speaker indifference implications are that the stronger propositions $[n]$ and $[n+1, \ldots)$ are not assertible because the speaker does not believe them to be true and she does not consider them to be relevant. Like the composite theory, the extended composite theory predicts that these implications are more robustly available with a superlative rather than a comparative modifier.

The predictions of the extended composite theory were tested by an extra Context condition (speaker indifference condition) included in our second eye-tracking experiment. We did not attest the same patterns for the indifference condition as for the ignorance condition, which was considered to be a challenge for the suitability of the extended composite theory. While no difference was attested between the two numeral modifier conditions at the region and measure where the relevant effects of ignorance had been observed (*ten people*, re-reading probability), there was even an indication that indifference is stronger with comparative modifiers (early effect at the numeral modifier region). Moreover, after having taken into consideration strong intuitions revealing that comparative modifiers are better cues to speaker indifference than

superlative modifiers are, I suggested that we could still maintain the extended composite theory as long as we took ignorance conventionalization into account for superlative modifiers. I explained that both types of numeral modifiers could trigger speaker ignorance and speaker indifference effects, as predicted by the extended composite theory, with the difference being that speaker ignorance with superlative modifiers is in the process of becoming conventionalized. This could result in comparative modifiers taking over the remaining function/implication, and, thus, appearing to be better at signaling speaker indifference in non-embedding environments. Lastly, I considered the possibility that the attested effects as far as the indifference condition was concerned could be the on-line manifestation of a different type of inference triggered by modified numerals in non-embedding environments with a knowledgeable and cooperative speaker, that is, the so-called granularity-based scalar implicatures (Cummins et al., 2012). Regardless of how we interpret the findings relating to the indifference Context condition of the second eye-tracking experiment, I concluded that we can maintain the extended composite theory taking conventionalization into account. Hence, this chapter came to complete the picture of the pragmatics of numeral modifiers built up throughout this thesis.

8.3 What has this thesis contributed

8.3.1 On a theoretical level

By looking at three different types of effects numeral modifiers can signal—variation, speaker ignorance, speaker indifference—this thesis achieved to contribute a more complete and global view of the pragmatics of numeral modifiers. A main finding is that superlative numeral modifiers give rise to non-obligatory, context-dependent specific ignorance implicatures. This finding helps to tell apart the numerous pragmatic accounts of ignorance effects of superlative modifiers, all being live options given the hitherto experimental findings. We took the relevant results to be at odds with Ciardelli et al. (2017); Cohen & Krifka (2014); Coppock & Brochhagen (2013b); Geurts & Nouwen (2007); Nouwen (2010); Spector (2015); Spychalska (2015), and consistent with the Quantity-based accounts by Büring (2008); Cummins & Katsos (2010); Kennedy (2015); Nouwen (2015), and Schwarz (2016a).

An equally significant finding is that the two types of numeral modifiers we focused on, namely, superlatives and comparatives, as representatives of class B and class A of modifiers respectively, trigger the same (set of) pragmatic inferences. This is in line with a few recent studies, such as Ciardelli et al. (2017); Mayr & Meyer (2014), and Westera & Brasoveanu (2014) (re speaker ignorance effects), and Mayr (2013) (re variation effects), whereas it invalidates a large number of studies, like Büring (2008); Coppock & Brochhagen (2013b); Cummins & Katsos (2010); Geurts & Nouwen (2007); Kennedy (2015); Nouwen (2010); Spychalska (2015) (re speaker ignorance) and Büring (2008); Coppock

& Brochhagen (2013b), and Nouwen (2015) (re variation). It was further shown that, although superlative and comparative modifiers trigger the same implicatures, they do so in varying degrees. Firstly, this suggests that Nouwen's (2010) well-established class A/B distinction, and as this was reformulated in Nouwen (2015), should not be seen as a classification of numeral modifiers as to the availability/absence of certain inferences, but rather as a distinction based on the likelihood or degree of the inferences in question.

This brings us to the next implication of the findings of this thesis. The observation as to the variability in robustness of pragmatic inferences is reminiscent of the finding by Doran et al. (2009, 2012), and van Tiel et al. (2016) that different scalar expressions give rise to scalar implicatures to a different degree. This observation has largely been ignored by formal theories. In this light, the composite theory put forth in this thesis could be viewed not only as an attempt to capture the defeasible/pragmatic, as opposed to semantic, nature of the inferences under investigation, but also as an attempt to capture the difference in robustness of those inferences between superlative and comparative numeral modifiers. Ciardelli et al. (2017) is another account along these lines, which nevertheless captures the difference in robustness between class A and class B modifiers only as far as ignorance inferences are concerned. Also, Ciardelli et al. derive a more robust, hard-to-cancel, and context-independent ignorance inference for superlative modifiers than that demonstrated by the overall findings of the present thesis.

But let us now refer back to my very initial question: Why does language provide us with various ways to express the same meaning? Or, taking into account what I have claimed so far, why would language provide us with an expression that triggers certain robust inferences (cf. superlatives) as well as with a truth-conditionally equivalent expression triggering the same inferences, only to a smaller degree (cf. comparatives)? Perhaps my suggestion to adopt the extended composite theory, but, crucially, including ignorance conventionalization considerations for superlative modifiers becomes relevant here. That is to say, such a proposal seems to give a satisfactory answer to the above question, as it points to a division of labor between superlative and comparative modified numerals with respect to the inferences they can trigger in non-embedding environments: Due to an underway conventionalization of ignorance with superlative modifiers, (strong) speaker ignorance effects are associated with superlative modifiers, and as a result comparative modifiers are better at signaling speaker indifference. Of course this remains to be clarified and verified by future research.

An attractive idea for further investigation would be to conduct a diachronic corpus study on numeral modifiers. Such a study would enable annotating and monitoring the various uses and effects of each numeral modifier (variation, ignorance, indifference) over sequential periods of time, allowing us to evaluate the aforementioned suggestion. As the robustness of ignorance implicatures is a cross-linguistic phenomenon (Nouwen, 2010), it would be interesting to carry out the corpus study in question in different languages. A large-scale

diachronic research along these lines has been conducted by Aloni, Aguilar-Guevara, Port, Simik, Solt, de Vos and Zeijlstra (see, e.g., Aloni, 2016) aiming to explore a number of different uses of free choice indefinites, such as free choice and speaker ignorance effects, in Spanish, Dutch, and German. Their research revealed that free choice inferences of the Spanish indefinite *cualquiera* and of the Dutch indefinite *wie dan ook* are derived as semantic entailments, while they are fossilized implicatures, resulting from semantic change, with the German indefinite *irgend/irgendein*. As to ignorance inferences, they concluded that those are the result of lexically encoded felicity conditions. As is obvious, diachronic corpus research can provide significant insight into the current status of different effects associated with certain expressions and, importantly, into how this has come about over time across different languages. Therefore, such a line of research can tell whether similar conclusions to Aloni's (2016) are to be drawn for the related variation and ignorance effects of numeral modifiers, or whether the proposed extended composite theory capturing both of them, and also indifference effects, as well as my supplementary diachronic suggestion are on the right track.

Given what has been concluded and discussed thus far, it should be clear that numeral modifiers trigger a variety of inferences, i.e., variation, speaker ignorance, and speaker indifference, and that we should look for a comprehensive theoretical framework to account for differences in likelihood/strength of inferences, such as those observed between superlative and comparative modifiers (with respect to the aforesaid inferences), possibly also taking into account conventionalization processes.

8.3.2 On a psycholinguistic level

The reading eye-tracking studies reported on in this thesis bring to light—direct or indirect—traces of the derivation of ignorance inferences, of indifference or granularity-based scalar inferences, and possibly also of a Manner-based reasoning. These findings are in accordance with other studies investigating Quantity- or informativity-related phenomena (e.g., Breheny et al., 2006; Engelhardt et al., 2006; Fukumura & van Gompel, 2017; Huang & Snedeker, 2009; Panizza et al., 2009), which also present evidence indicating that pragmatic reasoning is attestable in incremental interpretation. Our relevant findings further tend to suggest that stronger/more robust inferences occur earlier in incremental interpretation and are, specifically, more closely tied to the trigger, while weaker inferences might come up in a region much later than the trigger, cf. location differences of effects associated with the superlative vs. the comparative modifier.

Moreover, our on-line findings reveal that pragmatic reasoning manifests itself late in the eye movement record, i.e., in late measures such as the probability of re-reading and re-reading times, or in intermediate measures, like regression path duration and total reading times. This is in support of the idea that those measures are to be associated with higher-level processes.

Lastly, our findings go against a defaultist view on implicatures, like Levinson's (2000), requiring that implicatures are calculated by default and effortlessly. This is evident by the Context effects we observed in the eye-tracking experiments, showing that fewer/no implicatures were generated in the authority Context condition, as also supported by the high acceptability ratings the authority condition obtained in the relevant pretests. In fact, according to Levinson's (2000) account, the opposite Context effects should have been attested in the eye-tracking experiments due to the cancellation of the default ignorance interpretation in the authority contexts. On the other hand, our context-dependent effects are in line with the Context-Driven view on implicatures.

All in all, our study provides evidence from a new empirical domain that the calculation of pragmatic inferences is sensitive to contextual cues. To the extent that we can take the attested processing penalties to be directly caused by the on-line computation of speaker ignorance/indifference interpretations and/or of a Manner implicature, it is additionally suggested that such computations are effortful, lending further support to Context-Driven accounts of implicatures as presented, e.g., in Breheny et al. (2006) or Bott & Noveck (2004), where implicature computation comes with a processing cost. However, the gap between pragmatic theory and language use is still considerable. This thesis has shown that theory needs to be more fine-grained in order to connect to the data.

APPENDIX A

Experimental items of variation experiments

A.1 Experiment 1a

1.
Onderzoeker: Elk telefoonnummer in Afrika heeft **minstens/meer dan** vijf/ vijf **of meer** cijfers.
Interviewer: Hebben ze allemaal evenveel cijfers?

2.
Onderzoeker: In elke Nederlandse stad zijn er **minstens/meer dan** dertien/ dertien **of meer** raadsleden.
Interviewer: Zijn er evenveel raadsleden in elke Nederlandse stad?

3.
Onderzoeker: Elk gedicht van Hans Franke bevat **minstens/meer dan** vier/ vier **of meer** verborgen verwijzingen naar Dantes oeuvre.
Interviewer: Bevatten ze allemaal evenveel verborgen verwijzingen naar Dantes oeuvre?

4.
Onderzoeker: Tijdens het evenement werd elke straat door **minstens/meer dan** zes/ zes **of meer** agenten beveiligd.
Interviewer: Werden ze allemaal door evenveel agenten beveiligd?

5.
Onderzoeker: Elke grote supermarktketen heeft **minstens/meer dan** zestien/ zestien **of meer** vestigingen in de hoofdstad.
Interviewer: Hebben allemaal evenveel vestigingen in de hoofdstad?

6.
Onderzoeker: In elk restaurant in Enschede werken **minstens/meer dan** drie/ drie **of meer** obers.
Interviewer: Werken er in elk restaurant in Enschede evenveel obers?

A.2 Experiment 1b

1.
Onderzoeker: Elk telefoonnummer in Afrika heeft **maximaal/minder dan** elf/ elf **of minder** cijfers.
Interviewer: Hebben ze allemaal evenveel cijfers?

2.
Onderzoeker: In elke Nederlandse stad zijn er **maximaal/minder dan** twaalf/ twaalf **of minder** raadsleden.
Interviewer: Zijn er evenveel raadsleden in elke Nederlandse stad?

3.
Onderzoeker: Elk gedicht van Hans Franke bestaat uit **maximaal/minder dan** zes/ zes **of minder** strofes.
Interviewer: Bestaan ze allemaal uit evenveel strofes?

4.
Onderzoeker: Elke hoogleraar begeleidt **maximaal/minder dan** vier/ vier **of minder** bachelorscripties.
Interviewer: Begeleiden ze allemaal evenveel bachelorscripties?

5.
Onderzoeker: Elke grote supermarktketen heeft **maximaal/minder dan** vierentwintig/ vierentwintig **of minder** vestigingen in de hoofdstad.
Interviewer: Hebben ze allemaal evenveel vestigingen in de hoofdstad?

6.
Onderzoeker: Elk restaurant in Enschede heeft **maximaal/minder dan** negen/ negen **of minder** obers in dienst.
Interviewer: Hebben ze allemaal evenveel obers in dienst?

A.3 Experiment 1c

1.
Onderzoeker: Elk telefoonnummer in Afrika heeft **minstens/meer dan** vijf cijfers.
Test condition
Interviewer: Hebben ze allemaal evenveel cijfers?
Bad control
Interviewer: Zijn er geen met precies drie cijfers?
Good control
Interviewer: Hoe kwam je daarachter?

2.
Onderzoeker: In elke Nederlandse stad zijn er **minstens/meer dan** dertien raadsleden.
Test condition
Interviewer: Zijn er evenveel raadsleden in elke Nederlandse stad?
Bad control
Interviewer: Zijn er geen Nederlandse steden met precies tien raadsleden?
Good control
Interviewer: Hoe ben je daarachter gekomen?

3.
Onderzoeker: Elk gedicht van Hans Franke bevat **minstens/meer dan** vier verborgen verwijzingen naar Dantes oeuvre.
Test condition
Interviewer: Bevatten ze allemaal evenveel verborgen verwijzingen naar Dantes oeuvre?
Bad control
Interviewer: Zijn er geen met precies twee verborgen verwijzingen naar Dantes oeuvre?
Good control
Interviewer: Waar heb je dat gevonden?

4.
Onderzoeker: Tijdens het evenement werd elke straat door **minstens/meer dan** zes agenten beveiligd.
Test condition
Interviewer: Werden ze allemaal door evenveel agenten beveiligd?
Bad control
Interviewer: Was er geen straat met precies vier agenten?
Good control
Interviewer: Hoe weet je dat?

5.
Onderzoeker: Elke grote supermarktketen heeft **minstens/meer dan** zestien vestigingen in de hoofdstad.
Test condition
Interviewer: Hebben ze allemaal evenveel vestigingen in de hoofdstad?
Bad control
Interviewer: Zijn er geen supermarktketens met precies veertien winkels in Amsterdam?
Good control
Interviewer: Van wie heb je dat gehoord?

6.
Onderzoeker: In elk restaurant in Enschede werken **minstens/meer dan** drie obers.
Test condition
Interviewer: Werken er in elk restaurant in Enschede evenveel obers?
Bad control
Interviewer: Zijn er geen restaurants in Enschede waar precies twee obers werken?
Good control
Interviewer: Wie heeft je dat verteld?

7.
Onderzoeker: Elk boek van Jules Verne bevat **minstens/meer dan** zeven literaire figuren.
Test condition
Interviewer: Bevatten ze allemaal evenveel figuren?
Bad control
Interviewer: Zijn er geen boeken van Verne die precies vijf figuren hebben?
Good control
Interviewer: Waar heb je dat geleerd?

8.
Onderzoeker: Elke bankmanager in de onderzochte stad heeft **minstens/meer dan** vijftien medewerkers onder zich.
Test condition
Interviewer: Hebben de bankmanagers in de onderzochte stad allemaal evenveel medewerkers onder zich?
Bad control
Interviewer: Zijn er geen bankmanagers in de onderzochte stad met precies twaalf medewerkers onder zich?
Good control
Interviewer: Hoe weet je dat?

9.
Onderzoeker: Elke bloem in de Alpen heeft **minstens/meer dan** acht kroonbladeren.
Test condition
Interviewer: Hebben de bloemen in de Alpen allemaal evenveel kroonbladeren?
Bad control
Interviewer: Zijn er geen bloemen in de Alpen met precies zes kroonbladeren?
Good control
Interviewer: Waar heb je dat gehoord?

10.
Onderzoeker: Elke boer in Drenthe heeft **minstens/meer dan** acht koeien.
Test condition
Interviewer: Hebben de boeren in Drenthe allemaal evenveel koeien?
Bad control
Interviewer: Zijn er geen boeren in Drenthe die precies zes koeien hebben?
Good control
Interviewer: Hoezo denk je dat?

11.
Onderzoeker: Elk paar in de onderzochte gemeenschap had **minstens/meer dan** drie kinderen.
Test condition
Interviewer: Hadden ze allemaal evenveel kinderen?
Bad control
Interviewer: Was er geen met precies twee kinderen?
Good control
Interviewer: Hoe heb je dat gevonden?

12.
Onderzoeker: Elke huurder op deze postcode woont al **minstens/meer dan** negen jaar in zijn woning.
Test condition
Interviewer: Wonen ze allemaal evenveel jaar in hun woning?
Bad control
Interviewer: Zijn er geen die precies zeven jaar in hun huurwoning wonen?
Good control
Interviewer: Wie heeft je dat verteld?

A.4 Experiment 2a

1.
Onderzoeker: Alle telefoonnummers in Afrika hebben **minstens** vijf/**meer**

dan vier cijfers.
***Minimum value* test condition**
Interviewer: Waarom denk je dat er nummers met vijf cijfers zijn in Afrika?
***Higher value* test condition**
Interviewer: Waarom denk je dat er nummers met zes cijfers zijn in Afrika?

***Some* bad control condition**
Onderzoeker: Enkele telefoonnummers in Afrika hebben **minstens** vijf/**meer dan** vier cijfers.
Interviewer: Waarom denk je dat er geen nummers met meer dan vijf cijfers zijn in Afrika?

2.
Onderzoeker: In alle Nederlandse steden zijn er **minstens** dertien/**meer dan** twaalf raadsleden.
***Minimum value* test condition**
Interviewer: Wie heeft je verteld dat er een stad met dertien raadsleden is?
***Higher value* test condition**
Interviewer: Wie heeft je verteld dat er een stad met viertien raadsleden is?

***Some* bad control condition**
Onderzoeker: In enkele Nederlandse steden zijn er **minstens** dertien/**meer dan** twaalf raadsleden.
Interviewer: Wie heeft je verteld dat er geen stad met meer dan dertien raadsleden is?

3.
Onderzoeker: Alle gedichten van Hans Franke bevatten **minstens** vier/**meer dan** drie verborgen verwijzingen naar Dantes oevre.
***Minimum value* test condition**
Interviewer: Wie heeft je verteld dat er een gedicht met vier van die verwijzingen is?
***Higher value* test condition**
Interviewer: Wie heeft je verteld dat er een gedicht met vijf van die verwijzingen is?

***Some* bad control condition**
Onderzoeker: Enkele gedichten van Hans Franke bevatten **minstens** vier/**meer dan** drie verborgen verwijzingen naar Dantes oevre.
Interviewer: Wie heeft je verteld dat er geen gedichten met meer dan vier van die verwijzingen zijn?

4.
Onderzoeker: Tijdens het evenement werden alle straten door **minstens** vier/**meer dan** drie agenten beveiligd.

***Minimum value* test condition**
Interviewer: Hoe ben je er achtergekomen dat er een straat met vier agenten is?
***Higher value* test condition**
Interviewer: Hoe ben je er achtergekomen dat er een straat met vijf agenten is?

***Some* bad control condition**
Onderzoeker: Tijdens het evenement werden enkele straten door **minstens** vier/**meer dan** drie agenten beveiligd.
Interviewer: Hoe ben je er achtergekomen dat er geen straat met meer dan vier agenten was?

5.
Onderzoeker: Alle grote supermarktketens hebben **minstens** zestien/**meer dan** vijftien vestigingen in de hoofdstad.
***Minimum value* test condition**
Interviewer: Wie heeft je verteld dat er een supermarktketen is met zestien winkels in Amsterdam?
***Higher value* test condition**
Interviewer: Wie heeft je verteld dat er een supermarktketen is met zeventien winkels in Amsterdam?

***Some* bad control condition**
Onderzoeker: Enkele grote supermarktketens hebben **minstens** vijftien/**meer dan** vijftien vestigingen in de hoofdstad.
Interviewer: Wie heeft je verteld dat er geen supermarktketen is met meer dan zestien winkels in Amsterdam?

6.
Onderzoeker: Alle boeken van Jules Verne bevatten **minstens** zeven/**meer dan** zes literaire figuren.
***Minimum value* test condition**
Interviewer: Wie heeft je verteld dat er boeken van Verne zijn die zeven figuren hebben?
***Higher value* test condition**
Interviewer: Wie heeft je verteld dat er boeken van Verne zijn die acht figuren hebben?

***Some* bad control condition**
Onderzoeker: Enkele boeken van Jules Verne bevatten **minstens** zeven/**meer dan** zes literaire figuren.
Interviewer: Wie heeft je verteld dat er geen boeken van Verne zijn die meer dan zeven figuren hebben?

7.
Onderzoeker: Alle bankmanagers in de onderzochte stad hebben **minstens** drie/**meer dan** twee auto's.
***Minimum value* test condition**
Interviewer: Wie heeft je verteld dat er bankmanagers met drie auto's zijn?
***Higher value* test condition**
Interviewer: Wie heeft je verteld dat er bankmanagers met vier auto's zijn?

***Some* bad control condition**
Onderzoeker: Enkele bankmanagers in deze stad hebben **minstens** drie/**meer dan** twee auto's.
Interviewer: Wie heeft je verteld dat er geen bankmanagers met meer dan drie auto's zijn?

8.
Onderzoeker: Alle bloemen in de Alpen hebben **minstens** vier/**meer dan** drie kroonbladeren.
***Minimum value* test condition**
Interviewer: Waar heb je gevonden dat er bloemen in de Alpen zijn met vier kroonbladen?
***Higher value* test condition**
Interviewer: Waar heb je gevonden dat er bloemen in de Alpen zijn met vijf kroonbladen?

***Some* bad control condition**
Onderzoeker: Enkele bloemen in de Alpen hebben **minstens** vier/**meer dan** drie kroonbladeren.
Interviewer: Waar heb je gevonden dat er geen bloemen in de Alpen zijn met meer dan vier kroonbladen?

9.
Onderzoeker: In alle restaurants in Enschede werken **minstens** drie/**meer dan** twee obers.
***Minimum value* test condition**
Interviewer: Wie heeft je verteld dat er restaurants met drie obers zijn?
***Higher value* test condition**
Interviewer: Wie heeft je verteld dat er restaurants met vier obers zijn?

***Some* bad control condition**
Onderzoeker: In enkele restaurants in Enschede werken **minstens** drie/**meer dan** twee obers.
Interviewer: Wie heeft je verteld dat er geen restaurants met meer dan drie obers zijn?

10.
Onderzoeker: Alle boeren in Drenthe hebben **minstens** vijf/**meer dan** vier koeien.
***Minimum value* test condition**
Interviewer: Wie heeft je verteld dat er boeren zijn die vijf koeien hebben?
***Higher value* test condition**
Interviewer: Wie heeft je verteld dat er boeren zijn die zes koeien hebben?

***Some* bad control condition**
Onderzoeker: Enkele boeren in Drenthe hebben **minstens** vijf/**meer dan** vier koeien.
Interviewer: Wie heeft je verteld dat er geen boeren zijn die meer dan vijf koeien hebben?

11.
Onderzoeker: Alle paren in de onderzochte gemeenschap hadden **minstens** drie/**meer dan** twee kinderen.
***Minimum value* test condition**
Interviewer: Wie heeft je verteld dat er paren met drie kinderen waren?
***Higher value* test condition**
Interviewer: Wie heeft je verteld dat er paren met vier kinderen waren?

***Some* bad control condition**
Onderzoeker: Enkele paren in de onderzochte gemeenschap hadden **minstens** drie/**meer dan** twee kinderen.
Interviewer: Wie heeft je verteld dat er geen paren met meer dan drie kinderen waren?

12.
Onderzoeker: Alle huurders op deze postcode wonen al **minstens** drie/**meer dan** twee jaar in hun woning.
***Minimum value* test condition**
Interviewer: Hoe ben je er achtergekomen dat er bewoners zijn die drie jaar in hun huurwoning wonen?
***Higher value* test condition**
Interviewer: Hoe ben je er achtergekomen dat er bewoners zijn die vier jaar in hun huurwoning wonen?

***Some* bad control condition**
Onderzoeker: Enkele huurders op deze postcode wonen al **minstens** drie/**meer dan** twee jaar in hun woning.
Interviewer: Hoe ben je er achtergekomen dat er geen bewoners zijn die al meer dan drie jaar in hun huurwoning wonen?

13.
Onderzoeker: Alle waardes in het CERN-experiment van vorige week werden afgerond op **minstens** vier/**meer dan** drie cijfers achter de comma.
***Minimum value* test condition**
Interviewer: Hoe kwam je erachter dat er waardes waren die op vier cijfers achter de comma werden afgerond?
***Higher value* test condition**
Interviewer: Hoe kwam je erachter dat er waardes waren die op vijf cijfers achter de comma werden afgerond?

***Some* bad control condition**
Onderzoeker: Enkele waardes in het CERN-experiment van vorige week werden afgerond op **minstens** vier/**meer dan** drie cijfers achter de comma.
Interviewer: Hoe kwam je erachter dat er geen waardes waren die op meer dan vier cijfers achter de comma werden afgerond?

14.
Onderzoeker: In alle kooien in de dierentuin zitten **minstens** vier/**meer dan** drie dieren.
***Minimum value* test condition**
Interviewer: Wie heeft je verteld dat er kooien zijn met vier dieren?
***Higher value* test condition**
Interviewer: Wie heeft je verteld dat er kooien zijn met vijf dieren?

***Some* bad control condition**
Onderzoeker: In enkele kooien in de dierentuin zitten **minstens** vier/**meer dan** drie dieren.
Interviewer: Wie heeft je verteld dat er geen kooien zijn met meer dan vier dieren?

A.5 Experiment 2b

1.
Onderzoeker: Een universiteit moet **minstens** drie/**meer dan** twee verschillende computersystemen kopen.
***Minimum value* test condition**
Interviewer: Wie bepaalde dat het kopen van drie computersystemen genoeg is?
***Higher value* test condition**
Interviewer: Wie bepaalde dat het kopen van vier computersystemen genoeg is?

***Some* bad control condition**
Onderzoeker: Een universiteit mag **minstens** drie/**meer dan** twee verschillende computersystemen kopen.
Interviewer: Wie bepaalde dat het kopen van meer dan drie computersystemen teveel is?

2.
Onderzoeker: De piloot moet **minstens** vier/**meer dan** drie testvluchten met dit toestel maken.
***Minimum value* test condition**
Interviewer: Wie bepaalde dat het doen van vier vluchten genoeg is?
***Higher value* test condition**
Interviewer: Wie bepaalde dat het doen van vijf vluchten genoeg is?

***Some* bad control condition**
Onderzoeker: De piloot mag **minstens** vier/**meer dan** drie testvluchten met dit toestel maken.
Interviewer: Wie bepaalde dat het doen van meer dan vier vluchten teveel is?

3.
Onderzoeker: Een NASA-astronaut moet **minstens** vier/**meer dan** drie keer per maand oefenen met een vluchtsimulator.
***Minimum value* test condition**
Interviewer: Wie bepaalde dat vier keer per maand oefenen genoeg is?
***Higher value* test condition**
Interviewer: Wie bepaalde dat vijf keer per maand oefenen genoeg is?

***Some* bad control condition**
Onderzoeker: Een NASA-astronaut mag **minstens** vier/**meer dan** drie keer per maand oefenen met een vluchtsimulator.
Interviewer: Wie bepaalde dat meer dan vier keer per maand oefenen teveel is?

4.
Onderzoeker: Voor een artikel in deze krant moet een journalist **minstens** vijf/**meer dan** vier steekwoorden opgeven.
***Minimum value* test condition**
Interviewer: Wie heeft je verteld dat vijf steekwoorden genoeg is?
***Higher value* test condition**
Interviewer: Wie heeft je verteld dat zes steekwoorden genoeg is?

***Some* bad control condition**
Onderzoeker: Voor een artikel in deze krant mag een journalist **minstens** vijf/**meer dan** vier steekwoorden opgeven.
Interviewer: Wie heeft je verteld dat meer dan vijf steekwoorden teveel is?

5.
Onderzoeker: Een nieuwe burgemeester moet **minstens** zes/**meer dan** vijf externe adviseurs aanstellen.
***Minimum value* test condition**
Interviewer: Wanneer is er besloten dat het hebben van zes externe adviseurs genoeg is?
***Higher value* test condition**
Interviewer: Wanneer is er besloten dat het hebben van zeven externe adviseurs genoeg is?

***Some* bad control condition**
Onderzoeker: Een nieuwe burgemeester mag **minstens** zes/**meer dan** vijf externe adviseurs aanstellen.
Interviewer: Wanneer is er besloten dat het hebben van meer dan zes externe adviseurs teveel is?

6.
Onderzoeker: Volgens de richtlijnen moet een openbaar gebouw **minstens** vier/**meer dan** drie beveilingsmedewerkers hebben.
***Minimum value* test condition**
Interviewer: Wie bepaalde dat het aanstellen van vier beveilingsmedewerkers genoeg is?
***Higher value* test condition**
Interviewer: Wie bepaalde dat het aanstellen van vijf beveilingsmedewerkers genoeg is?

***Some* bad control condition**
Onderzoeker: Volgens de richtlijnen mag een openbaar gebouw **minstens** vier/**meer dan** drie beveilingsmedewerkers hebben.
Interviewer: Wie bepaalde dat het aanstellen van meer dan vier beveilingsmedewerkers teveel is?

7.
Onderzoeker: Bij de zitting moeten **minstens** zestien/**meer dan** vijftien commissieleden aanwezig zijn.
***Minimum value* test condition**
Interviewer: Wanneer is er besloten dat de aanwezigheid van zestien commissieleden genoeg is?
***Higher value* test condition**
Interviewer: Wanneer is er besloten dat de aanwezigheid van zeventien commissieleden genoeg is?

***Some* bad control condition**
Onderzoeker: Bij de zitting mogen **minstens** zestien/**meer dan** vijftien com-

missieleden aanwezig zijn.
Interviewer: Wanneer is er besloten dat de aanwezigheid van meer dan zestien commissieleden teveel is?

8.
Onderzoeker: In een Intercity-trein moeten **minstens** drie/**meer dan** twee conducteurs aanwezig zijn.
***Minimum value* test condition**
Interviewer: Sinds wanneer zijn er richtlijnen dat de aanwezigheid van drie conducteurs genoeg is?
***Higher value* test condition**
Interviewer: Sinds wanneer zijn er richtlijnen dat de aanwezigheid van vier conducteurs genoeg is?

***Some* bad control condition**
Onderzoeker: In een Intercity-trein mogen **minstens** drie/**meer dan** twee conducteurs aanwezig zijn.
Interviewer: Sinds wanneer zijn er richtlijnen dat de aanwezigheid van meer dan drie conducteurs teveel is?

9.
Onderzoeker: Een militaire oefening moet door **minstens** vijf/**meer dan** vier officieren gesurveilleerd worden.
***Minimum value* test condition**
Interviewer: Wie bepaalde dat de aanwezigheid van vijf officieren genoeg is?
***Higher value* test condition**
Interviewer: Wie bepaalde dat de aanwezigheid van zes officieren genoeg is?

***Some* bad control condition**
Onderzoeker: Een militaire oefening mag door **minstens** vijf/**meer dan** vier officieren gesurveilleerd worden.
Interviewer: Wie bepaalde dat de aanwezigheid van meer dan vijf officieren teveel is?

10.
Onderzoeker: Een Grieks bankrekeningnummer moet **ten minste** zes/**meer dan** vijf cijfers hebben.
***Minimum value* test condition**
Interviewer: Wanneer is er besloten dat het hebben van zes cijfers genoeg is?
***Higher value* test condition**
Interviewer: Wanneer is er besloten dat het hebben van zeven cijfers genoeg is?

***Some* bad control condition**
Onderzoeker: Een Grieks bankrekeningnummer mag **ten minste** zes/**meer**

dan vijf cijfers hebben.
Interviewer: Wanneer is er besloten dat het hebben van meer dan zes cijfers teveel is?

11.
Onderzoeker: Een productlijn van Nokia moet **minstens** drie/**meer dan** twee telefoons tegelijk ontwikkelen.
***Minimum value* test condition**
Interviewer: Wie bepaalde dat het ontwikkelen van drie telefoons tegelijk genoeg is?
***Higher value* test condition**
Interviewer: Wie bepaalde dat het ontwikkelen van vier telefoons tegelijk genoeg is?

***Some* bad control condition**
Onderzoeker: Een productlijn van Nokia mag **minstens** drie/**meer dan** twee telefoons tegelijk ontwikkelen.
Interviewer: Wie bepaalde dat het ontwikkelen van meer dan drie telefoons tegelijk teveel is?

12.
Onderzoeker: In een kathedraal moeten er **minstens** drie/**meer dan** twee priesters in dienst zijn.
***Minimum value* test condition**
Interviewer: Wie bepaalde dat het in dienst hebben van drie priesters genoeg is?
***Higher value* test condition**
Interviewer: Wie bepaalde dat het in dienst hebben van vier priesters genoeg is?

***Some* bad control condition**
Onderzoeker: In een kathedraal mogen er **minstens** drie/**meer dan** twee priesters in dienst zijn.
Interviewer: Wie bepaalde dat het in dienst hebben van meer dan drie priesters teveel is?

13.
Onderzoeker: Een verzekeringsagent moet **minstens** negen/**meer dan** acht klanten hebben.
***Minimum value* test condition**
Interviewer: Wie bepaalde dat het hebben van negen klanten genoeg is?
***Higher value* test condition**
Interviewer: Wie bepaalde dat het hebben van tien klanten genoeg is?

***Some* bad control condition**
Onderzoeker: Een verzekeringsagent mag **minstens** negen/**meer dan** acht klanten hebben.
Interviewer: Wie bepaalde dat het hebben van meer dan negen klanten teveel is?

14.
Onderzoeker: Voor de directeur van Phillips moeten **minstens** vier/**meer dan** drie secretarissen werken.
***Minimum value* test condition**
Interviewer: Wie heeft je verteld dat het hebben van vier secretarissen genoeg is?
***Higher value* test condition**
Interviewer: Wie heeft je verteld dat het hebben van vijf secretarissen genoeg is?

***Some* bad control condition**
Onderzoeker: Voor de directeur van Phillips mogen **minstens** vier/**meer dan** drie secretarissen werken.
Interviewer: Wie heeft je verteld dat het hebben van meer dan vier secretarissen teveel is?

APPENDIX B

Experimental items of eye-tracking experiment 1

1.
Sophie is een kunstschaatsster en erg fanatiek. Afgelopen weekend ging ze proberen zo intensief mogelijk te trainen.
Ignorance Context condition
Ik weet niet helemaal zeker hoeveel exact, maar dit is mijn idee:
Authority Context condition
Ik kan je melden hoeveel omdat ik gisteren met haar gepraat heb.
Modal Verb condition
Sophie wilde minstens zeven uur op het ijs oefenen.
Non-modal Verb condition
Sophie heeft minstens zeven uur op het ijs geoefend.
Ze is al weken bezig met een erg lastige nieuwe oefening.

2.
Hans is een bouwvakker en hij heeft een grote klus binnengehaald. Hij is constant bezig met het slopen en opbouwen van schuren.
Ignorance Context condition
Het is niet helemaal duidelijk hoeveel, maar ik zou je mijn schatting kunnen geven.
Authority Context condition
Hij sprak er met me over en daarom kan ik je er precies over vertellen.
Modal Verb condition
Hans moest minstens twee schuren op woensdagmiddag afbreken.
Non-modal Verb condition
Hans heeft minstens twee schuren op woensdagmiddag afgebroken.

Hij vindt het sloopwerk het leukst omdat hij hierbij niet zo voorzichtig hoeft te zijn.

3.
Hugo was een dag niet op zijn werk geweest en zijn secretaresse kon niet al zijn telefoongesprekken beantwoorden. Toen Hugo dinsdagochtend weer op zijn werk verscheen moest hij daarom meteen hard werken.
Ignorance Context condition
Ik weet niet helemaal exact hoe het zit, maar ik zal je mijn voorstelling geven.
Authority Context condition
Ik weet exact hoe het zit omdat hij hierover tegen mij klaagde.
Modal Verb condition
Hij moest minstens twintig telefoontjes in de ochtend afhandelen.
Non-modal Verb condition
Hij heeft minstens twintig telefoontjes in de ochtend afgehandeld.

4.
Jim is pakketbezorger en hij heeft het tegenwoordig ook in de avonden erg druk. Hij rijdt heel Eindhoven op en neer om pakjes te bezorgen.
Ignorance Context condition
Hij heeft me geen informatie gegeven over hoe druk hij het had, maar het zal ongeveer zo zijn:
Authority Context condition
Hij heeft me informatie gegeven over hoe druk hij het had en het zit zo:
Modal Verb condition
Jim moest minstens vierentwintig pakketten in de avond afleveren.
Non-modal Verb condition
Jim heeft minstens vierentwintig pakketten in de avond afgeleverd.
In de avonden krijgt Jim extra betaald en daarom vindt hij het niet erg dat hij het tegenwoordig zo druk heeft in de avonden.

5.
De studente Merel had veel deadlines voor haar studie biologie. Ze had voor zichzelf een planning gemaakt omdat ze zoveel opdrachten moest maken dat ze door de bomen het bos niet meer zag.
Ignorance Context condition
Het is voor mij niet helemaal duidelijk, maar het zal ongeveer zoiets zijn geweest:
Authority Context condition
Ze had het er gisteren over en daarom is het duidelijk voor mij.
Modal Verb condition
Merel wilde minstens drie opdrachten op maandag afmaken.
Non-modal Verb condition
Merel heeft minstens drie opdrachten op maandag afgemaakt.
Voordat Merel aan haar studie begon had ze niet verwacht dat ze zo vaak op-

drachten zou moeten inleveren.

6.
Linda probeerde haar favoriete spijkerbroek aan te trekken, maar deze ging opeens moeizaam dicht. Ze nam zich daarom voor om een aantal kilo kwijt te raken.
Ignorance Context condition
Ik weet niet exact hoe het zit, maar ik zal je een schatting geven.
Authority Context condition
Door een gesprek met Linda ben ik volkomen op de hoogte.
Modal Verb condition
Linda wilde minstens twee en een halve kilo in een maand afvallen.
Non-modal Verb condition
Linda is minstens twee en een halve kilo in een maand afgevallen.
Ze heeft altijd al erg moeite gehad om op hetzelfde gezonde gewicht te blijven.

7.
De 85-jarige mevrouw De Jong zit in een verzorgingstehuis en kijkt er altijd erg naar uit als de kleinkinderen op bezoek komen. Ze vindt het leuk om ze cadeaus te geven en het liefst iets wat ze zelf heeft gebreid.
Ignorance Context condition
Ik weet niet precies hoe het zit, maar ik zal je vertellen hoe ik denk dat het is.
Authority Context condition
Ik weet precies hoe het zit, omdat ze hierover enthousiast met me heeft gekletst.
Modal Verb condition
Mevrouw De Jong wilde minstens vier truien in de winter breien.
Non-modal Verb condition
Mevrouw De Jong heeft minstens vier truien in de winter gebreid.

8.
Meneer Jansen werkt bij de onderwijsinspectie en erzijn afgelopen week verschillende rapporten over scholen geschreven. Hij was van plan alle rapporten deze week door te nemen en dat zorgde voor een grote drukte.
Ignorance Context condition
Ik ben niet op de hoogte van de precieze drukte, maar het zal hierbij in de buurt zijn geweest:
Authority Context condition
Hij informeerde mij gisteren over de drukte en daarom kan ik je het volgende zeggen:
Modal Verb condition
Meneer Jansen moest minstens negen rapporten op dinsdag doornemen.
Non-modal Verb condition
Meneer Jansen heeft minstens negen rapporten op dinsdag doorgenomen.

9.
Charlotte is met haar familie op vakantie naar Kenia. Ze houdt erg veel van fotograferen en dit wilde ze dan ook erg graag doen tijdens de safari's.
Ignorance Context condition
Ik ben niet helemaal op de hoogte, maar ik zal je vertellen wat ik denk.
Authority Context condition
Ik ben helemaal op de hoogte, omdat ze zit dit ter sprake bracht.
Modal Verb condition
Ze wilde minstens vijftien diersoorten op vakantie fotograferen.
Non-modal Verb condition
Ze heeft minstens vijftien diersoorten op vakantie gefotografeerd.
Volgend jaar wil ze naar China om daar fotos te maken van authentieke gebouwen.

10.
De studenten die het vak wetenschapsfilosofie volgen moesten een artikel schrijven. Ze vonden de opdracht alleen erg lastig en hadden daarom veel vragen voor de docent. De docent wilde de studenten zo goed mogelijk helpen en plande hier veel tijd voor in.
Ignorance Context condition
Helemaal precies weet ik niet hoe het zit, maar ik heb wel een vaag idee.
Authority Context condition
We hebben gisteren even gepraat en daarom kan ik je erover inlichten.
Modal Verb condition
De docent moest minstens elf studenten op woensdag helpen.
Non-modal Verb condition
De docent heeft minstens elf studenten op woensdag geholpen.
De studenten waren in ieder geval gemotiveerd.

11.
Eriks ouders hadden vroeger nooit heel veel geld en ze kochten al hun kleding bij de kringloopwinkel. Daarom wilde Erik nu de kringloopwinkels helpen door ze kleding aan te bieden. Hij startte een grote inzamelingsactie in Urk.
Ignorance Context condition
Zijn concept was me niet helemaal duidelijk, maar ik zal je vertellen wat ik denk.
Authority Context condition
Ik heb het er met hem over gehad en daarom zal ik je erover vertellen.
Modal Verb condition
Erik wilde minstens honderdvijftig kledingstukken in Urk inzamelen.
Non-modal Verb condition
Erik heeft minstens honderdvijftig kledingstukken in Urk ingezameld.
Zijn ouders zijn hierdoor erg trots op Erik.

12.
Alexander heeft een goedlopend Grieks restaurant en op vrijdagavond zijn al-

tijd alle tafels gereserveerd. Er wordt van de kok verlangd dat hij hard werkt en veel gerechten bereidt.
Ignorance Context condition
Hij heeft dit niet echt aan me verduidelijkt, maar ik zal je beschrijven hoe ik denk dat het zal zijn.
Authority Context condition
We hebben het vanochtend besproken en ik kan je het daarom beschrijven.
Modal Verb condition
De kok moest minstens vierennegentig gerechten op vrijdagavond klaarmaken.
Non-modal Verb condition
De kok heeft minstens vierennegentig gerechten op vrijdagavond klaargemaakt.
De volgende dag kon hij gelukkig wel lekker uitslapen.

13.
An is secretaresse bij een bank en moet elke dag veel documenten kopiëren. Maandagmiddag zou er een belangrijke vergadering zijn en An moest ervoor zorgen dat iedereen de juiste documenten zou krijgen.
Ignorance Context condition
Ze heeft het me niet echt toegelicht, maar ik zal je mijn schatting geven.
Authority Context condition
Ze bracht me op de hoogte van de klus en daarom kan ik je het volgende zeggen:
Modal Verb condition
An moest minstens vierhonderd pagina's op maandagochtend kopiëren.
Non-modal Verb condition
An heeft minstens vierhonderd pagina's op maandagochtend gekopieerd.
Soms vindt ze haar werk wel een beetje saai.

14.
In juli gaat Moniek voor een maand naar Spanje en daarom is ze nu hard bezig om Spaans te leren. De grammatica gaat al aardig, maar ze wil haar woordenschat nog meer uitbreiden.
Ignorance Context condition
Het is voor mij wat vaag, maar ik zal je mijn idee geven.
Authority Context condition
Het is voor mij duidelijk en daarom zal ik je erover vertellen.
Modal Verb condition
Moniek wilde minstens veertig werkwoorden op maandag leren.
Non-modal Verb condition
Moniek heeft minstens veertig werkwoorden op maandag geleerd.
In Spanje wil ze ook heel graag de beginselen van de flamenco leren.

15.
Op vrijdagavond fiets ik altijd langs een boerderij in de buurt en maak ik een praatje met de boer. Hij heeft veel koeien die allemaal elke ochtend gemolken moeten worden.

Ignorance Context condition
Ik weet niet precies hoe het zit, maar ik zal je mijn schatting geven.
Authority Context condition
Ik weet precies hoe het zit, omdat hij er uitgebreid over sprak.
Modal Verb condition
De boer moest minstens vijfenvijftig koeien voor het ontbijt melken.
Non-modal Verb condition
De boer heeft minstens vijfenvijftig koeien voor het ontbijt gemolken.

16.
Ad is docent wiskunde op een middelbare school. Zijn leerlingen hadden vrijdagmiddag een proefwerk gemaakt en hij wilde het zo snel mogelijk nagekeken hebben. Hij zou daarom op zijn vrije zaterdagmiddag alvast beginnen.
Ignorance Context condition
Het is mij niet helemaal duidelijk, maar ik geef je mijn globale idee.
Authority Context condition
Het is mij helemaal duidelijk en ik kan daarom dit opmerken:
Modal Verb condition
Ad wilde minstens twaalf proefwerken op zijn vrije middag nakijken.
Non-modal Verb condition
Ad heeft minstens twaalf proefwerken op zijn vrije middag nagekeken.
Bij zijn leerlingen is Ad geliefd omdat hij proefwerken altijd zo snel mogelijk nakijkt.

17.
Kunstenaar Floris tekent het liefst de hele dag door. Hij maakt vaak portretten van mensen, maar ook dieren vindt hij erg leuk om te tekenen. Hij ging vandaag naar de dierentuin om een aantal tekeningen te maken.
Ignorance Context condition
Ik heb niet echt inzicht in de situatie, maar ik zal je vertellen hoe ik vermoed dat het zit.
Authority Context condition
We hebben gisteren even gepraat en daarom kan ik je dit melden:
Modal Verb condition
Floris wilde minstens vijf dieren in de dierentuin natekenen.
Non-modal Verb condition
Floris heeft minstens vijf dieren in de dierentuin nagetekend.
Hij weet ondertussen precies de weg in de dierentuin, omdat hij er zo vaak komt.

18.
Thomas en Fleur hebben een nieuw huis gekocht met een gigantische achtertuin. Ze hebben geen tijd om de achtertuin zelf aan te pakken en daarom hebben ze een tuinman ingehuurd. Hij kreeg van Thomas en Fleur de vrije hand bij het kappen en planten van bomen.

Ignorance Context condition
Ik ben er niet echt over ingelicht, maar ik kan erover speculeren.
Authority Context condition
Hij praatte vol trots over deze klus en daarom kan ik je hiervan verslag uitbrengen.
Modal Verb condition
De tuinman wilde minstens dertien bomen in de achtertuin omkappen.
Non-modal Verb condition
De tuinman heeft minstens dertien bomen in de achtertuin omgekapt.

19.
Op het station in Den Haag zit een bloemist waar het altijd erg druk is. De grootste drukte is tijdens het middaguur en daarom wilde de bloemist al hard doorwerken in de ochtend.
Ignorance Context condition
Ik kan je niet exact vertellen hoe het zit, maar ik zal je mijn schatting geven.
Authority Context condition
Ik kan je exact vertellen hoe het zit, omdat hij mij dit had toevertrouwd.
Modal Verb condition
De bloemist wilde minstens zeventien boeketten in de ochtend samenstellen.
Non-modal Verb condition
De bloemist heeft minstens zeventien boeketten in de ochtend samengesteld.
Er worden naast boeketten ook altijd veel vazen verkocht.

20.
Vivian werkt als schoonmaakster bij een hotel. Ze wil alle kamers altijd grondig schoonmaken, omdat ze graag wil dat de gasten tevreden zijn.
Ignorance Context condition
Het was me niet helemaal helder, maar dit is mijn voorstelling:
Authority Context condition
Het was me helemaal helder en daarom kan ik je dit zeggen:
Modal Verb condition
Vivian moest minstens veertien kamers tijdens haar werkdag schoonmaken.
Non-modal Verb condition
Vivian heeft minstens veertien kamers tijdens haar werkdag schoongemaakt.
Haar baas waardeert haar ijver erg.

21.
Mia is kok bij een groot restaurant en vandaag stond er als voorgerecht een salade op het menu. Ze had de tomaat al gesneden en daarna ging ze met de komkommers aan de gang. Ze wilde erg opschieten, want het restaurant ging al bijna open.
Ignorance Context condition
Ik weet niet precies hoe het zit, maar ik zal aan je melden wat ik vermoed.

Authority Context condition
Ze vertelde me erover en daarom weet ik precies hoe het zit.
Modal Verb condition
Mia moest minstens vierenzestig komkommers in een kwartier snijden.
Non-modal Verb condition
Mia heeft minstens vierenzestig komkommers in een kwartier gesneden.
Ze vindt het nooit erg om hard te werken, omdat ze altijd veel complimenten krijgt.

22.
Mevrouw Zwart kreeg haar hele familie op bezoek. Ze wilde graag het huis aan kant hebben en daarom had ze voor zichzelf een planning gemaakt. Ze begon in de ochtend met stofzuigen.
Ignorance Context condition
Het idee was voor mij niet helemaal duidelijk, maar ik zal je over mijn gedachte vertellen.
Authority Context condition
Het idee was voor mij helemaal duidelijk, omdat ze dit luid verkondigde.
Modal Verb condition
Mevrouw Zwart wilde minstens acht kamers in de ochtend stofzuigen.
Non-modal Verb condition
Mevrouw Zwart heeft minstens acht kamers in de ochtend gestofzuigd.
Ze vindt schoonmaken geen vervelende taak en is altijd al erg netjes geweest.

23.
Wesley heeft zijn eigen zaak waar hij met veel plezier tatoeages zet. Hij maakt er altijd echte kunstwerken van en het is er dan ook meestal erg druk. Hij wilde deze donderdag daarom hard doorwerken.
Ignorance Context condition
Ik weet niet precies hoe het met de drukte zat, maar ik heb wel een idee.
Authority Context condition
Ik weet precies hoe het met de drukte zat en daarom zal ik je erover vertellen.
Modal Verb condition
Wesley moest minstens zes mensen op donderdag tatoeëren.
Non-modal Verb condition
Wesley heeft minstens zes mensen op donderdag getatoeëerd.
Hij heeft ook twee medewerkers in dienst die de kunst van het tatoeren van hem hebben geleerd.

24.
Stijn heeft zijn eigen bakkerswinkel en verkoopt daar brood en gebak. Woensdagochtend had hij veel broden gebakken en zijn hele winkel stond er vol mee.
Ignorance Context condition
Ik weet niet precies hoe het zit, maar ik zal je een schatting geven.

Authority Context condition
Ik weet precies hoe het zit en daarom zal ik je het volgende vertellen:
Modal Verb condition
Stijn moest minstens vijftig broden op woensdag verkopen.
Non-modal Verb condition
Stijn heeft minstens vijftig broden op woensdag verkocht.
Zijn zus helpt hem altijd mee met het verkopen van brood en gebak.

25.
In juni is er in Zwolle een zwemwedstrijd en Ralph wil hier graag aan meedoen. Hij heeft echter weinig ervaring met wedstrijdzwemmen en is daarom hard aan het trainen geslagen.
Ignorance Context condition
Het is voor mij niet helemaal inzichtelijk, maar ik heb wel een idee.
Authority Context condition
Hij heeft me er vanochtend over geïnformeerd en daarom zal ik het je beschrijven.
Modal Verb condition
Ralph wilde minstens dertig baantjes op zaterdag zwemmen.
Non-modal Verb condition
Ralph heeft minstens dertig baantjes op zaterdag gezwommen.
Hij wist toen hij begon met trainen niet hoeveel techniek er bij wedstrijdzwemmen komt kijken.

26.
Kaspar is als accountant ingehuurd om de financile situatie van een bedrijf te analyseren. Hij sprak met de financieel directeur over de omzet van het bedrijf.
Ignorance Context condition
Ik weet niet wat er precies is gezegd, maar ik heb een vaag idee.
Authority Context condition
Ik sprak met Kasper erover en daarom zal ik je hier nu over inlichten.
Modal Verb condition
De financieel directeur moest minstens honderddertig miljoen euro in het laatje binnenkrijgen.
Non-modal Verb condition
De financieel directeur heeft minstens honderddertig miljoen euro in het laatje binnenkrijgen.
Volgend jaar moet de omzet nog meer omhoog.

27.
Nina houdt van sporten in de buitenlucht en wilde gaan trainen voor een marathon. Het was prachtig weer en daarom ging ze naar het bos.
Ignorance Context condition
Het is me niet helemaal helder, maar ik zal je mijn indruk geven.

Authority Context condition
We kletsten hierover en daarom kan ik je dit melden:
Modal Verb condition
Nina wilde minstens zesentwintig kilometer in het bos hardlopen.
Non-modal Verb condition
Nina heeft minstens zesentwintig kilometer in het bos hardgelopen.
Haar vriend is trots dat Nina zo fanatiek traint.

28.
Voor het goede doel wilde Lars proberen zoveel mogelijk rondjes rondom het meer voor zijn huis te fietsen. Hij had hier een uur voor en hij was gesponsord door vrienden en familie. Voor elk rondje zou hij een bepaald bedrag krijgen en daarom wilde hij natuurlijk zo goed mogelijk zijn best doen.
Ignorance Context condition
Ik ben niet geheel op de hoogte, maar ik zal je vertellen wat ik denk.
Authority Context condition
Lars heeft minstens achtentwintig rondjes om het meer gereden.
Modal Verb condition
Lars wilde minstens achtentwintig rondjes om het meer rijden.
Non-modal Verb condition
Lars heeft minstens achtentwintig rondjes om het meer gereden.

29.
Sven wil graag professioneel zanger worden en daarom organiseert hij een concert. Hij heeft een bandje uit zijn klas gevraagd om op te treden en Sven zelf zal ook wat nummers zingen. Het hoe en wat is me niet helemaal duidelijk, maar ik heb wel een vermoeden.
Ignorance Context condition
Het hoe en wat is me niet helemaal duidelijk, maar ik heb wel een vermoeden.
Authority Context condition
Hij vertelde me erover en daarom kan ik je hierover informeren.
Modal Verb condition
Sven moest minstens tweehonderd kaartjes voor het concert verkopen.
Non-modal Verb condition
Sven heeft minstens tweehonderd kaartjes voor het concert verkocht.
Het concert zal plaatsvinden in de kantine van zijn middelbare school.

30.
Roos had een tussenjaar genomen tijdens haar studie en ze was van plan te gaan reizen en werken. Ze wilde graag voor een tijdje in Canada rondtrekken.
Ignorance Context condition
Ze heeft me heel weinig over haar reis verteld, maar ik zal je vertellen wat ik vermoed.
Authority Context condition
We hebben uitgebreid over haar reis gesproken en daarom kan ik deze kennis

met je delen.
Modal Verb condition
Roos wilde minstens vijfenveertig dagen in Canada vertoeven.
Non-modal Verb condition
Roos heeft minstens vijfenveertig dagen in Canada vertoefd.
Haar ouders hebben ook altijd veel gereisd, dus het is geen wonder dat Roos reizen ook leuk vindt.

31.
Meneer Willemsen heeft in Nederland een aantal goedlopende winkels waar hij kookspullen verkoopt. Vorig jaar heeft hij een stel winkels geopend in Duitsland en hij hoopte dat deze het ook erg goed zouden doen.
Ignorance Context condition
Ik kan niet precies aangeven hoe goed, maar ik zal je een ingeving van me geven.
Authority Context condition
Ik sprak er met hem over en daarom zal ik je er nu over berichten.
Modal Verb condition
Meneer Willemsen moest minstens zes en een half miljoen euro winst voor het einde van vorig jaar maken.
Non-modal Verb condition
Meneer Willemsen heeft minstens zes en een half miljoen euro winst voor het einde van vorig jaar gemaakt.
Voor dit jaar verwacht hij een stijging van de winst.

32.
Professor Vermeulen is bezig met een psychologisch onderzoek waarbij ze naar emoties van mannen en vrouwen kijkt. Hiervoor heeft ze een experiment ontwikkeld dat ze zou afnemen onder studenten.
Ignorance Context condition
Ze heeft het niet aan me uitgelegd, maar ik zal je mijn gedachte toevertrouwen.
Authority Context condition
Ze heeft het aan me uitgelegd en daarom zal ik je een beschrijving geven.
Modal Verb condition
Professor Vermeulen wilde minstens vijfenzestig proefpersonen bij haar onderzoek betrekken.
Non-modal Verb condition
Professor Vermeulen heeft minstens vijfenzestig proefpersonen bij haar onderzoek betrokken.

APPENDIX C

Experimental items of eye-tracking experiment 2

1.
Sophie is een kunstschaatsster en erg fanatiek. Ze traint normaal vier uur in een weekend, maar afgelopen weekend heeft ze zo intensief mogelijk getraind.
Ignorance Context condition
Ik weet niet helemaal zeker hoeveel exact, maar dit is mijn idee:
Authority Context condition
Ik kan je melden hoeveel omdat ik gisteren met haar gepraat heb.
Indifference Context condition
Ik zou je exact kunnen vertellen hoeveel, maar dat is niet zo belangrijk.
Sophie heeft het afgelopen weekend **minstens/meer dan** acht uur op het ijs geoefend.
Ze is al weken bezig met een erg lastige nieuwe oefening.

2.
Hans is een bouwvakker en constant bezig met het slopen en opbouwen van schuren. Op een gewone werkdag zou hij slechts n schuur slopen of opbouwen, maar woensdag sloopte hij er meer.
Ignorance Context condition
Het is niet helemaal duidelijk hoeveel, maar ik zou je mijn schatting kunnen geven.
Authority Context condition
Hij sprak er met me over en daarom kan ik je er precies over vertellen.
Indifference Context condition
Hij sprak er met me over, maar ik zal er verder niet in detail over vertellen.
Hans heeft op woensdagmiddag **minstens/meer dan** vijf schuren met behulp

van zijn sloopmachine afgebroken.
Hij vindt het sloopwerk het leukst omdat hij hierbij niet zo voorzichtig hoeft te zijn.

3.
Op een drukke werkdag handelt Hugo gewoonlijk vijftien telefoontjes af. Hij was een dag niet op zijn werk geweest en toen hij dinsdagochtend weer op zijn werk verscheen moest hij daarom meteen hard werken.
Ignorance Context condition
Ik weet niet helemaal exact hoe het zit, maar ik zal je mijn voorstelling geven.
Authority Context condition
Ik weet exact hoe het zit, omdat hij hierover tegen mij klaagde.
Indifference Context condition
Ik weet exact hoe het zit, maar zo interessant is dat niet om te weten.
Hugo heeft in de ochtend **minstens/meer dan** twintig telefoontjes zo snel mogelijk afgehandeld.

4.
Jim is pakketbezorger in Eindhoven en hij heeft het tegenwoordig ook in de avonden erg druk. Hij bezorgde vandaag overdag twintig pakketjes.
Ignorance Context condition
Hij heeft me weinig informatie gegeven over hoe druk hij het had, maar het zal ongeveer zo zijn:
Authority Context condition
Hij heeft me informatie gegeven over hoe druk hij het had en het zit zo:
Indifference Context condition
Hij heeft me informatie gegeven over hoe druk hij het had, maar daar zal ik nu niet in detail op ingaan:
Jim heeft 's avonds **minstens/meer dan** vijfentwintig pakketten aan klanten afgeleverd.
In de avonden krijgt Jim extra betaald en daarom vindt hij het niet erg dat hij het tegenwoordig zo druk heeft in de avonden.

5.
Linda probeerde een tijd geleden haar favoriete spijkerbroek aan te trekken, maar deze ging opeens moeizaam dicht. Daarom wilde ze een aantal kilo afvallen. Eerder dit jaar was ze al eens vijf kilo kwijtgeraakt, maar die kilo's waren er snel weer bijgekomen.
Ignorance Context condition
Ik weet niet exact hoe het zit, maar ik zal je een schatting geven.
Authority Context condition
Door een gesprek met Linda ben ik volkomen op de hoogte.
Indifference Context condition
Ik heb met Linda gesproken, maar het is niet van belang dat je alles exact weet.
Linda is de afgelopen maand **minstens/meer dan** twee kilos met veel afzien

afgevallen.

6.
De 85-jarige mevrouw De Jong breidt in de winter vaak een trui voor zichzelf. Ze vindt het leuk om haar kleinkinderen zelfgemaakte cadeaus te geven, en daarom wilde ze voor de kerst nu ook een trui voor een aantal van haar kleinkinderen breien.
Ignorance Context condition
Ik weet niet precies hoe het zit, maar ik zal je vertellen hoe ik denk dat het is.
Authority Context condition
Ik weet precies hoe het zit, omdat ze hierover enthousiast met me heeft gekletst.
Indifference Context condition
Ik weet precies hoe het zit, maar het is niet zo belangrijk dat je even goed op de hoogte bent als ik.
Mevrouw De Jong heeft voor kerstmis **minstens/meer dan** vier truien als cadeau gebreid.

7.
Meneer Jansen werkt bij de onderwijsinspectie en er zijn afgelopen week verschillende rapporten over scholen geschreven. Dat zorgde voor grote drukte. Maandag nam hij alvast de eerste twee rapporten door.
Ignorance Context condition
Ik ben niet op de hoogte van de precieze drukte, maar het zal hierbij in de buurt zijn geweest:
Authority Context condition
Hij informeerde mij gisteren over de drukte en daarom kan ik je het volgende zeggen:
Indifference Context condition
Hij informeerde mij gisteren over de drukte, maar ik zal niet te diep in gaan op de details.
Meneer Jansen heeft op dinsdag **minstens/meer dan** tien rapporten over scholen doorgenomen.

8.
Loes is met haar familie op vakantie naar Kenia, waar ze veel foto's wilde maken tijdens de safari's. Haar vader, die ook van foto- graferen houdt, fotografeerde vijf verschillende diersoorten.
Ignorance Context condition
Ik ben niet helemaal op de hoogte, maar ik zal je vertellen wat ik denk.
Authority Context condition
Ik ben helemaal op de hoogte, omdat ze dit ter sprake bracht.
Indifference Context condition
Ik ben helemaal op de hoogte, maar ik vertel je nu niet alles wat ik weet.
Loes heeft tijdens haar vakantie **minstens/meer dan** vijftien diersoorten met

haar splinternieuwe camera gefotografeerd.
Volgend jaar wil ze naar China om daar fotos te maken van authentieke gebouwen.

9.
De studenten die het vak wetenschapsfilosofie volgen moesten een artikel schrijven. Ze vonden de opdracht erg lastig, maar de docent wilde de studenten graag helpen en plande hier veel tijd voor in. Vorig jaar hoefde hij maar drie studenten te helpen.
Ignorance Context condition
Helemaal precies weet ik niet hoe het zit, maar ik heb wel een vaag idee.
Authority Context condition
We hebben gisteren even gepraat en daarom kan ik je erover inlichten.
Indifference Context condition
We hebben gis- teren even gepraat, maar ik ga je nu niet heel nauwkeurig inlichten.
De docent heeft woensdag **minstens/meer dan** tien studenten met het werk geholpen.
De studenten waren in ieder geval gemotiveerd.

10.
Eriks ouders kochten vroeger al hun kleding bij de kringloopwinkel. Daarom wilde Erik nu iets terug doen voor deze winkels. In Emmeloord had hij al vijftig kledingstukken ingezameld met een kleinschalige inzamelingsactie. Nu startte hij een grote inzamelingsactie in Urk.
Ignorance Context condition
De opbrengst is me niet helemaal duidelijk, maar ik zal je vertellen wat ik denk.
Authority Context condition
Ik heb het er met hem over gehad en daarom zal ik je erover vertellen.
Indifference Context condition
Ik heb het er met hem over gehad, maar ik zal niet alles exact herhalen.
Erik heeft de afgelopen week **minstens/meer dan** honderdvijftig kledingstukken in Urk ingezameld.
Zijn ouders zijn erg trots op hem.

11.
Alexander heeft een goedlopend Grieks restaurant. De kok heeft deze week gemiddeld veertig gerechten per avond klaargemaakt, maar vrijdagavond was het nog drukker.
Ignorance Context condition
Hij heeft dit niet echt aan me verduidelijkt, maar ik zal je beschrijven hoe ik denk dat het zal zijn.
Authority Context condition
We hebben het vanochtend besproken en ik kan je het daarom beschrijven.
Indifference Context condition
We hebben het vanochtend besproken, maar ik zal het alleen in grote lijnen

beschrijven.
De kok heeft op vrijdagavond **minstens/meer dan** zestig gerechten voor gasten klaargemaakt.
De volgende dag kon hij gelukkig wel lekker uitslapen.

12.
An is secretaresse bij een bank en moet elke dag veel documenten kopiren. Ze kopieert op de meeste dagen in totaal driehonderd pagina's. Maandagmiddag zou er een belangrijke vergadering zijn.
Ignorance Context condition
Ze heeft het me niet echt toegelicht, maar ik zal je mijn schatting geven.
Authority Context condition
Ze bracht me op de hoogte van de klus en daarom kan ik je het volgende zeggen:
Indifference Context condition
Ze bracht me op de hoogte van de klus, maar de details vertel ik nu niet.
An heeft op maandagochtend **minstens/meer dan** vierhonderd pagina's voor haar collega's gekopieerd.
Soms vindt ze haar werk wel een beetje saai.

13.
In juli gaat Moniek samen met haar vriend voor een maand naar Spanje en daarom is ze nu hard bezig om Spaans te leren. Deze week probeert ze ongeveer dertig werkwoorden per dag te leren.
Ignorance Context condition
Het is voor mij wat vaag, maar ik zal je mijn idee geven.
Authority Context condition
Het is voor mij duidelijk en daarom zal ik je erover vertellen.
Indifference Context condition
Het is voor mij duidelijk, maar ik ga er niet al te diep op in.
Moniek heeft op maandag **minstens/meer dan** veertig werkwoorden met hulp van Haar vriend geleerd.
In Spanje wil ze ook heel graag de beginselen van de flamenco leren.

14.
Ik maak vaak een praatje met de boer waar ik verse melk koop. Vandaag sprak ik hem weer. Hij heeft honderdveertig koeien die elke ochtend gemolken moeten worden.
Ignorance Context condition
Ik weet niet precies hoe het zit, maar ik zal je mijn schatting geven.
Authority Context condition
Ik weet precies hoe het zit, omdat hij er uitgebreid over sprak.
Indifference Context condition
Ik weet precies hoe het zit, maar zal het niet uitgebreid navertellen.
De boer heeft voor het ontbijt **minstens/meer dan** vijftig koeien met behulp van zijn melkmachine gemolken.

15.
De leerlingen van Ad hebben vrijdagmiddag een proefwerk gemaakt. Hij wilde het snel nagekeken hebben en daarom zou hij op zijn vrije zaterdagmiddag alvast beginnen. Vaak krijgt hij het voor elkaar om zestig proefwerken na te kijken op één dag.
Ignorance Context condition
Het is mij niet helemaal duidelijk, maar ik geef je mijn globale idee.
Authority Context condition
Het is mij helemaal duidelijk en ik kan daarom dit opmerken:
Indifference Context condition
Het is mij helemaal duidelijk, maar ik ga je niet exact melden wat hij me vertelde.
Ad heeft op zijn vrije middag **minstens/meer dan** veertig proefwerken in snel tempo nagekeken.
Bij zijn leerlingen is Ad geliefd omdat hij proefwerken altijd snel nakijkt.

16.
Kunstenaar Floris maakt portretten van mensen, maar ook vaak van dieren. Hij gaat graag naar de dierentuin en tekent dan meestal een of twee dieren na. Dit weekend ging hij ook weer naar de dierentuin, nu om zo veel mogelijk tekeningen te maken.
Ignorance Context condition
Ik heb niet echt inzicht in de situatie, maar ik zal je vertellen hoe ik vermoed dat het zit.
Authority Context condition
We hebben gisteren even gepraat en daarom kan ik je dit melden:
Indifference Context condition
We hebben gisteren even gepraat, maar ik wil je nu niet in alle detail inlichten.
Floris heeft dit weekend **minstens/meer dan** vijf dieren in de dierentuin nagetekend.
Hij weet ondertussen precies de weg in de dierentuin, omdat hij er zo vaak komt.

17.
Thomas en Fleur hebben een nieuw huis gekocht met een gigantische achtertuin waar veel bomen in staan. Ze hebben geen tijd om de achtertuin zelf aan te pakken en daarom hebben ze een tuinman ingehuurd. De tuinman is van plan tien bomen te kappen en is daar vanochtend mee begonnen.
Ignorance Context condition
Ik ben er niet echt over ingelicht, maar ik kan erover speculeren.
Authority Context condition
Hij praatte vol trots over deze klus en daarom kan ik je hiervan verslag uitbrengen.
Indifference Context condition
Hij praatte vol trots over deze klus, maar ik zal alleen in grote lijnen erover

vertellen.
De tuinman heeft vanochtend **minstens/meer dan** vijftien bomen in de achtertuin omgekapt.

18.
Op het station in Den Haag zit een bloemist waar het altijd erg druk is, vooral tijdens het middaguur. Hij verkoopt dan meestal twintig boeketten. Daarom wilde de bloemist deze ochtend weer hard doorwerken.
Ignorance Context condition
Ik kan je niet exact vertellen hoe het zit, maar ik zal je mijn schatting geven.
Authority Context condition
Ik kan je exact vertellen hoe het zit, omdat hij mij dit had toevertrouwd.
Indifference Context condition
Ik kan je exact vertellen hoe het zit, maar ik ga het nu niet grondig beschrijven.
De bloemist heeft in de ochtend **minstens/meer dan** vijfentwintig boeketten met veel zorg samengesteld.
Er worden naast boeketten ook altijd veel vazen verkocht.

19.
Vivian werkt als schoonmaakster bij een hotel. Op een normale werkdag maakt ze vijftien kamers schoon, maar vandaag werkte ze extra hard door.
Ignorance Context condition
Het is me niet helemaal helder, maar dit is mijn voorstelling:
Authority Context condition
Het is me helemaal helder en daarom kan ik je dit zeggen:
Indifference Context condition
Het is me helemaal helder, maar ik breng je geen uitgebreid verslag.
Vivian heeft vandaag **minstens/meer dan** twintig kamers tijdens haar werkdag schoongemaakt.
Haar baas waardeert haar ijver erg.

20.
Mia is kok bij een groot restaurant en was bezig met het snijden van komkommers. Ze wilde erg opschieten, want het restaurant ging al bijna open. Normaal gesproken doet ze een kwartier over het snijden van vijftig komkommers.
Ignorance Context condition
Ik weet niet precies hoe het zit, maar ik zal aan je melden wat ik vermoed.
Authority Context condition
Ze vertelde me erover en daarom weet ik precies hoe het zit.
Indifference Context condition
Ze heeft me er alles over verteld, maar dat ga ik nu niet exact herhalen.
Mia heeft vandaag **minstens/meer dan** zeventig komkommers in een kwartier gesneden.
Ze vindt het nooit erg om hard te werken, omdat ze altijd veel complimenten krijgt.

21.
Mevrouw Zwart kreeg haar hele familie op bezoek. Ze wilde graag het huis aan kant hebben en begon in de ochtend met stofzuigen. Gisteren had ze toevallig al twee kamers gestofzuigd, dus die kon ze overslaan.
Ignorance Context condition
Het idee was voor mij niet helemaal duidelijk, maar ik zal je over mijn gedachte vertellen.
Authority Context condition
Het idee was voor mij helemaal duidelijk, omdat ze dit luid verkondigde.
Indifference Context condition
Het idee was voor mij helemaal duidelijk, maar je hoeft het niet exact te weten.
Mevrouw Zwart heeft 's ochtends **minstens/meer dan** vier kamers met grote precisie gestofzuigd.
Ze vindt schoonmaken geen vervelende taak en is altijd al erg netjes geweest.

22.
Wesley heeft zijn eigen zaak waar hij met veel plezier tatoeages zet. Het is er meestal erg druk en hij probeert elke dag acht mensen te tatoeëren. Deze donderdag was hij ook weer hard aan het werk.
Ignorance Context condition
Ik weet niet precies hoe het met de drukte zat, maar ik heb wel een idee.
Authority Context condition
Ik weet precies hoe het met de drukte zat en daarom zal ik je erover vertellen.
Indifference Context condition
Ik weet precies hoe het met de drukte zat, maar dat is niet zo interessant om uitvoerig na te vertellen.
Wesley heeft die dag **minstens/meer dan** tien mensen met veel oog voor detail getatoeëerd.
Hij heeft ook twee medewerkers in dienst die de kunst van het tatoeëren van hem hebben geleerd.

23.
Stijn heeft zijn eigen bakkerswinkel en verkoopt daar brood en gebak. Hij verkoopt doordeweeks gemiddeld veertig broden per dag, maar woensdag was er een speciale korting op het brood voor zijn vaste klanten.
Ignorance Context condition
Ik weet niet precies hoe het zit, maar ik zal je een schatting geven.
Authority Context condition
Ik weet precies hoe het zit en daarom zal ik je het volgende vertellen:
Indifference Context condition
Ik weet precies hoe het zit, maar zal er niet uitvoerig op in gaan:
Stijn heeft op woensdag **minstens/meer dan** vijftig broden met korting verkocht.
Zijn zus helpt hem altijd mee met het verkopen van brood en gebak.

24.
In juni is er een zwemwedstrijd en Ralph wil hier graag aan meedoen. Hij is daarom hard aan het trainen geslagen. Voordat hij begon met trainen, kon hij twintig baantjes achter elkaar zwemmen.
Ignorance Context condition
Het is voor mij niet helemaal inzichtelijk, maar ik heb wel een idee.
Authority Context condition
Hij heeft me er vanochtend over geïnformeerd en daarom zal ik het je beschrijven.
Indifference Context condition
Ik heeft me er vanochtend over genformeerd, maar het is niet zo relevant om alles precies na te vertellen.
Ralph heeft op zaterdag **minstens/meer dan** dertig baantjes achter elkaar gezwommen.
Hij wist toen hij begon met trainen niet hoeveel techniek er bij wedstrijdzwemmen komt kijken.

25.
Nina houdt van sporten in de buitenlucht en traint nu voor een marathon. Het was gisteren prachtig weer en daarom ging ze naar het bos. Drie dagen geleden liep ze een rondje van twintig kilometer.
Ignorance Context condition
Het is me niet helemaal helder, maar ik zal je mijn indruk geven.
Authority Context condition
We kletsten hierover en daarom kan ik je dit melden:
Indifference Context condition
We kletsten hierover, maar de details zijn nu niet zo relevant.
Nina heeft gisteren **minstens/meer dan** vijfentwintig kilometer in het bos hardgelopen.
Haar vriend is trots dat Nina zo fanatiek traint.

26.
Voor het goede doel wilde Lars gisteren proberen binnen een uur zoveel mogelijk rondjes rondom het meer voor zijn huis te fietsen. Voor elk rondje zou hij een bepaald bedrag krijgen en daarom wilde hij natuurlijk zo goed mogelijk zijn best doen. Hij hoopte twintig rondjes te kunnen fietsen.
Ignorance Context condition
Ik ben niet geheel op de hoogte, maar ik zal je vertellen wat ik denk.
Authority Context condition
Ik heb er met Lars over gepraat en daarom kan ik je hierover vertellen.
Indifference Context condition
Ik heb er met Lars over gepraat, maar ik zal je slechts een globaal idee geven.
Lars heeft uiteindelijk **minstens/meer dan** dertig rondjes om het meer gereden.

27.
Sven is zanger in een band en zij treden vrijdagavond op. Sven heeft zich toegelegd op de verkoop van kaartjes. Er kunnen driehonderd kaartjes verkocht worden.
Ignorance Context condition
Het hoe en wat is me niet helemaal duidelijk, maar ik heb wel een vermoeden.
Authority Context condition
Hij vertelde me erover en daarom kan ik je hierover informeren.
Indifference Context condition
Hij vertelde me erover, maar ik zal je nu alleen een globaal idee geven.
Sven heeft de afgelopen week **minstens/meer dan** tweehonderd kaartjes voor het concert verkocht.
Het concert zal plaatsvinden in de kantine van zijn middelbare school.

28.
Roos had een tussenjaar genomen tijdens haar studie en wilde graag in Canada rondtrekken. Ze had genoeg gespaard om hopelijk twee maanden rond te kunnen reizen.
Ignorance Context condition
Ze heeft me heel weinig over haar reis verteld, maar ik zal je vertellen wat ik vermoed.
Authority Context condition
We hebben uitgebreid over haar reis gesproken en daarom kan ik deze kennis met je delen.
Indifference Context condition
We hebben uitgebreid over haar reis gesproken, maar ik zal het niet uitvoerig herhalen.
Roos heeft tijdens haar tussenjaar **minstens/meer dan** vijfenveertig dagen in Canada vertoefd.
Haar ouders hebben ook altijd veel gereisd, dus het is geen wonder dat Roos reizen ook leuk vindt.

29.
Meneer Willemsen heeft in Nederland een aantal goedlopende winkels. Vorig jaar heeft hij een winkel geopend in Duitsland en het eerste jaar heeft hij meteen al een half miljoen euro winst gemaakt. Hij hoopte dat de winst verder zou stijgen.
Ignorance Context condition
Ik kan niet precies aangeven hoe goed, maar ik zal je een ingeving van me geven.
Authority Context condition
Ik sprak er met hem over en daarom zal ik je er nu over berichten.
Indifference Context condition
Ik sprak er met hem over, maar de details zijn momenteel niet zo interessant.
Meneer Willemsen heeft dit jaar **minstens/meer dan** twee miljoen euro winst

met zijn nieuwe winkel gemaakt.
Volgend jaar verwacht hij weer een stijging van de winst.

30.
Bart is afgelopen zondag naar een groter huis verhuisd. Tijdens zijn vorige verhuizing hoefde hij nog niet zoveel spullen te verhuizen. Hij verhuisde toen van een studentenkamer naar een klein appartement en hoefde slechts tien dozen in- en uit te pakken. Deze keer was dat anders.
Ignorance Context condition
Het is mij niet helemaal duidelijk, maar ik kan je wel een ingeving geven.
Authority Context condition
Hij heeft me hierover ingelicht en daarom kan ik je hierover informatie geven.
Indifference Context condition
Hij heeft me er uitvoerig over verteld, maar ik houd het simpel:
Bart heeft op zondag **minstens/meer dan** twintig dozen een voor een uitgepakt.

31.
Vandaag was er een groot atletiektoernooi waar Tobias aan meedeed. Hij hoopte vooral een medaille te kunnen behalen met het discuswerpen. Zijn persoonlijk record staat op vijfentwintig meter, maar daarmee liep hij tijdens het vorige toernooi net de bronzen medaille mis.
Ignorance Context condition
Ik weet niet helemaal hoe het zit, maar ik zal je mijn schatting geven.
Authority Context condition
Ik weet helemaal hoe het zit, dus dat zal ik nu aan je bekendmaken.
Indifference Context condition
Ik weet helemaal hoe het zit, maar ik zal het niet te uitvoerig beschrijven.
Tobias heeft vandaag de discus **minstens/meer dan** dertig meter het veld in geworpen.

32.
Femke wilde graag een bijbaantje zodat ze een extra zakcentje zou hebben. Ze werkt daarom nu als vakkenvuller bij de plaatselijke supermarkt. Ze heeft een contract voor twaalf uur per week, maar soms maakt ze extra uren.
Ignorance Context condition
Ze heeft me hier amper informatie over gegeven, maar ik zal je vertellen wat ik denk.
Authority Context condition
Ze heeft me hier uitgebreid over geïnformeerd en daarom kan ik je hierover inlichten.
Indifference Context condition
Ze heeft me goed geïnformeerd, maar ik ga het nu niet uitgebreid navertellen.
Femke heeft de afgelopen week **minstens/meer dan** vijftien uur in de supermarkt gewerkt.

33.
Suus ging vroeger elk jaar samen met haar ouders op vakantie naar Frankrijk. Ze heeft op die manier het land goed leren kennen en besloot dat ze er graag een tijdje zou willen wonen. Ze was van plan drie jaar te blijven en dan weer terug te verhuizen naar Nederland.
Ignorance Context condition
Ik heb er niet veel inzicht in, maar ik zal je een vermoeden van me vertellen.
Authority Context condition
We hebben er uitgebreid over gepraat, dus ik kan je erover inlichten.
Indifference Context condition
We hebben er uitgebreid over gepraat, maar ik zal nu niet teveel vertellen.
Suus heeft uiteindelijk **minstens/meer dan** vier jaar in Frankrijk gewoond.

34.
Victor is dit jaar begonnen met vioollessen. Hij doet erg zijn best en probeert elke dag een half uur te oefenen. Ook op zondag oefent hij viool. Hij wil namelijk zo snel mogelijk in een orkest kunnen meespelen.
Ignorance Context condition
Het is me niet helemaal duidelijk, maar ik zal je vertellen wat ik denk.
Authority Context condition
Het is me helemaal duidelijk en daarom kan ik je het volgende mededelen:
Indifference Context condition
Het is me helemaal duidelijk, maar details ga ik je nu niet geven.
Victor heeft zondag **minstens/meer dan** twee uur op zijn viool gespeeld.

35.
Jet studeert geschiedenis in Amsterdam. Ze is bezig met het schrij- ven van haar scriptie die aan het eind van de maand af moet zijn. Ze moet in totaal tienduizend woorden schrijven.
Ignorance Context condition
Ik weet niet exact hoe het zit, maar het zal ongeveer zo zijn:
Authority Context condition
Ik heb hierover met haar gepraat en daarom kan ik je er in detail over inlichten.
Indifference Context condition
Ik heb hierover met haar gepraat, maar zal je nu geen overbodige informatie geven.
Jet heeft de afgelopen twee weken **minstens/meer dan** drieduizend woorden op papier gezet.

36.
Anouk is een bekend schrijfster. Twintig jaar geleden publiceerde ze haar eerste boek en sinds die tijd zijn er veel boeken bijgekomen. De eerste tien jaar heeft ze voornamelijk detectives geschreven. Dat waren er in totaal zeven. Daarna is ze ook boeken binnen andere genres gaan schrijven.

Ignorance Context condition
Ik ben niet helemaal op de hoogte, maar ik kan erover speculeren.
Authority Context condition
Ik las vanochtend een artikel over haar, dus ik kan je er alles over vertellen.
Indifference Context condition
Ik las vanochtend een artikel over haar, maar ik ga je nu niet volledig inlichten.
Anouk heeft het afgelopen decennium **minstens/meer dan** vijf romans vlot achter elkaar geschreven.

Bibliography

Albrecht, J. E. & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1061–1070.

Alexandropoulou, S. (2015). Testing the nature of variation effects with modified numerals. In Csipak, E. & Zeijlstra, H. (Eds.), *Sinn und Bedeutung*, volume 19, (pp. 36–53).

Alexandropoulou, S., Dotlačil, J., McNabb, Y., & Nouwen, R. (2015). Pragmatic inferences with numeral modifiers: Novel experimental data. In D'Antonio, S., Moroney, M., & Little, C. R. (Eds.), *Semantics and Linguistic Theory (SALT)*, volume 25, (pp. 533–549).

Alexandropoulou, S., Dotlačil, J., & Nouwen, R. (2016). *At least* ignorance inferences come at a processing cost: Support from eye movements. In Moroney, M., Little, C.-R., Collard, J., & Burgdorf, D. (Eds.), *Semantics and Linguistic Theory (SALT)*, volume 26, (pp. 795–813).

Alexandropoulou, S., Dotlačil, J., & Nouwen, R. (2017). Pragmatic effects of *more than* and *at least* in incremental interpretation. In Burgdorf, D. & Collard, J. (Eds.), *Semantics and Linguistic Theory (SALT)*, volume 27.

Aloni, M. (2007). Expression ignorance or indifference. Modal implicatures in BiOT. In *Proceedings of the Sixth International Tbilisi Symbisium on Language, Logic and Computation*, (pp. 1–20).

Aloni, M. (2016). Indefinites as fossils. Paper presented at 'Indefinites between Theory and Language Change' workshop (DGfS), University of Konstanz, February 24.

Aloni, M. & van Rooij, R. (2007). Free Choice Items and Alternatives. In G. Bouma, I. Kraemer, & J. Zwarts (Eds.), *Cognitive Foundations of Interpretation* (pp. 333–352). Amsterdam: Edita KNAW.

Alonso-Ovalle, L. & Menéndez-Benito, P. (2010). Modal indefinites. *Natural Language Semantics*, *18*(1), 1–31.

Alonso-Ovalle, L. & Menéndez-Benito, P. (2011). Expressing indifference: Spanish *un* NP *cualquiera*. In Ashton, N., Chereches, A., & Lutz, D. (Eds.), *Semantics and Linguistic Theory (SALT)*, volume 21, (pp. 333–352).

Alonso-Ovalle, L. & Menéndez-Benito, P. (2013). Two views on epistemic indefinites. *Language and Linguistics Compass*, *7*(2), 105–122.

Anderson, C. (2004). *The structure and real-time comprehension of quantifier scope ambiguity.* PhD thesis, Northwestern University, Chicago.

Ariel, M. (2004). Most. *Language*, *80*(4), 658–706.

Asher, N. (2012). The non cooperative basis of implicatures. In *Logical Aspects of Computational Linguistics*, volume 7351, (pp. 45–57).

Asher, N. & Lascarides, A. (2013). Strategic conversation. *Semantics & Pragmatics*, *6*(2), 1–58.

Bach, K. (2006). The top 10 misconceptions about implicature. In B. Birner & G. Ward (Eds.), *Drawing the boundaries of meaning: Neo-Gricean studies in pragmatics and semantics in honor of Laurence R. Horn* (pp. 21–30). Amsterdam/Philadelphia: John Benjamins.

Baker, L. & Anderson, R. I. (1982). Effects of inconsistent information on text processing: Evidence from comprehension monitoring. *Reading Research Quarterly*, *17*(2), 281–294.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Barwise, J. & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, *4*(2), 159–219.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bergen, L. & Grodner, D. J. (2012). Speaker knowledge influences the comprehension of pragmatic inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1450–1460.

Blok, D. (2018). Phd thesis. Utrecht University.

Bohan, J. & Sanford, A. (2008). Semantic anomalies at the borderline of consciousness: An eye-tracking investigation. *Quarterly Journal of Experimental Psychology*, *61*(2), 232–239.

Boland, J. E. (2004). Linking eye movements to sentence comprehension in reading and listening. In M. Carreiras & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond* (pp. 51–76). New York/Hove: Psychology Press.

Boston, M. F., Hale, J. T., Vasishth, S., & Kliegl, R. (2011). Parallel processing and sentence comprehension difficulty. *Language and Cognitive Processes*, *26*(3), 301–349.

Bott, L., Bailey, T. M., & Grodner, D. (2012). Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language*, *66*(1), 123–142.

Bott, L., Frisson, S., & Murphy, G. L. (2009). Interpreting conjunctions. *The Quarterly Journal of Experimental Psychology*, *62*(4), 681–706.

Bott, L. & Noveck, I. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language*, *51*(3), 437–457.

Breheny, R., Ferguson, H. J., & Katsos, N. (2013). Investigating the time-course of accessing conversational implicatures during incremental sentence interpretation. *Language and Cognitive Processes*, *28*(4), 443–467.

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition*, *100*(3), 434–463.

Buccola, B. & Haida, A. (2017). A surface-scope analysis of authoritative readings of modified numerals. In *Sinn und Bedeutung*, volume 21.

Büring, D. (2008). The least *at least* can do. In Chang, C. B. & Haynie, H. J. (Eds.), *West Coast Conference on Formal Linguistics*, volume 26, (pp. 114–120).

Chemla, E. & Singh, R. (2014a). Remarks on the experimental turn in the study of scalar implicature, Part I. *Language and Linguistics Compass*, *8*(9), 373–386.

Chemla, E. & Singh, R. (2014b). Remarks on the experimental turn in the study of scalar implicature, Part II. *Language and Linguistics Compass*, *8*(9), 387–399.

Chevalier, C., Noveck, I., Nazir, T., Bott, L., Lanzetti, V., & Sperber, D. (2008). Making disjunctions exclusive. *Quarterly Journal of Experimental Psychology*, *61*, 1750–1761.

Chevallier, C., Bonnefond, M., van der Henst, J.-B., & Noveck, I. A. (2010). Using ERP's to capture inferential processes guided by prosodic cues. *Italian Journal of Linguistics*, *22*(1), 125–152.

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the "locality" of language. *Linguistic Inquiry*, *37*(4), 535–590.

Chierchia, G. (2013). *Logic in grammar: Polarity, free choice, and intervention.* Oxford: Oxford University Press.

Chierchia, G., Fox, D., & Spector, B. (2013). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In P. Portner, C. Maienborn, & K. von Heusinger (Eds.), *An international handbook of natural language meaning* (pp. 2297–2332). Berlin: De Gruyter Mouton.

Choi, J. (2005). Another Type of Free Choice Effect: Korean *Amwu* N-*Lato.* In John Alderete, C.-h. H. & Kochetov, A. (Eds.), *West Coast Conference on Formal Linguistics*, volume 24.

Choi, J. (2007). *Free choice and negative polarity: A compositional analysis of Korean polarity sensitive items.* PhD thesis, University of Pennsylvania, Pennsylvania.

Christensen, R. H. B. (2013). Ordinal: Regression models for ordinal data. [R package version 2013.9-30].

Ciardelli, I., Coppock, L., & Roelofsen, F. (2017). Implicatures of modified numerals: Quality or Quantity? In *Sinn und Bedeutung*, volume 21.

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. van Gompel, M. Fisher, W. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain*, volume 27 (pp. 341–372). Amsterdam: Elsevier.

Cohen, A. & Krifka, M. (2010). Superlative quantifiers as modifiers of meta-speech acts. In M. Glanzberg, B. Partee, & J. Škilters (Eds.), *The Baltic international yearbook of cognition, logic and communication*, volume 6 (pp. 1–56). Kansas: New Prairie Press.

Cohen, A. & Krifka, M. (2014). Superlative quantifiers and meta-speech acts. *Linguistics and Philosophy*, *37*(1), 41–90.

Condoravdi, C. (2015). Ignorance, indifference, and individuation in wh-*ever.* In L. Alonso-Ovalle & P. Benito-Ménendez (Eds.), *Epistemic indefinites: Exploring modality beyond the verbal domain* (pp. 213–243). Oxford: Oxford University Press.

Coppock, E. & Brochhagen, T. (2013a). Diagnosing truth, interactive sincerity, and depictive sincerity. In Snider, T., Jackson, B., & Matthews, T. (Eds.), *Semantics and Linguistic Theory*, volume 23, (pp. 358–375).

Coppock, E. & Brochhagen, T. (2013b). Raising and resolving issues with scalar modifiers. *Semantics & Pragmatics*, *6*(3), 1–57.

Cozijn, R. (2006). Het gebruik van oogbewegingen in leesonderzoek [The use of eye movements in reading research]. *Tijdschrift voor Taalbeheersing*, *28*, 220–232.

Crnič, L., Chemla, E., & Fox, D. (2015). Scalar implicatures of embedded disjunction. *Natural Language Semantics*, *23*(4), 271–305.

Cummins, C. & Katsos, N. (2010). Comparative and superlative quantifiers: Pragmatic effects of comparison type. *Journal of Semantics*, *27*(3), 271–305.

Cummins, C., Sauerland, U., & Solt, S. (2012). Granularity and scalar implicature in numerical expressions. *Linguistics and Philosophy*, *35*, 135–169.

Degen, J. & Tanenhaus, M. (2014). Processing scalar implicature: A constraint-based approach. *Cognitive Science*, *39*(4), 667–710.

Degen, J. & Tanenhaus, M. K. (2011). Making inferences: The case of scalar implicature processing. In Carlson, L., Hölscher, C., & Shipley, T. (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, (pp. 3299–3304).

Doran, R., Baker, R. E., McNabb, Y., Larson, M., & Ward, G. (2009). On the non-unified nature of scalar implicature: An empirical investigation. *International Review of Pragmatics*, *1*(2), 211–248.

Doran, R., Ward, G., Larson, M., McNabb, Y., & Baker, R. E. (2012). A novel experimental paradigm for distinguishing between 'what is said' and 'what is implicated'. *Language*, *88*(1), 124–154.

Drummond, A. (2007). Ibex, (version 0.3.8) [experiment platform]. http://spellout.net/ibexfarm/.

Engelhardt, P., Bailey, K., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean maxim of quantity? *Journal of Memory and Language*, *54*(4), 554–573.

Foppolo, F. & Marelli, M. (2017). No delay for some inferences. *Journal of Semantics*, *34*(4), 659–681.

Fox, D. (2014). Cancelling the maxim of quantity: Another challenge for a Gricean theory of scalar implicatures. *Semantics & Pragmatics*, *7*(5), 1–20.

Fox, D. & Hackl, M. (2006). The universal density of measurement. *Linguistics and Philosophy*, *29*(5), 537–586.

Fukumura, K. & van Gompel, R. P. (2017). How do violations of Gricean maxims affect reading? *Journal of Memory and Language*, *95*, 1–18.

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

Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* Cambridge: Cambridge University Press.

Geurts, B. (2011). *Quantity Implicatures.* Cambridge: Cambridge University Press.

Geurts, B., Katsos, N., Cummins, C., Moons, J., & Noordman, L. (2010). Scalar quantifiers: Logic, acquisition and processing. *Language and Cognitive Processes*, *25*(1), 130–48.

Geurts, B. & Nouwen, R. (2007). At least et al.: the semantics of scalar modifiers. *Language*, *83*(3), 533–559.

Geurts, B. & Rubio-Fernández, P. (2015). Pragmatics and Processing. *Ratio*, *28*(4), 446–469.

Gibson, E., Piantadosi, S. T., & Levy, R. (2017). Post hoc analysis decisions drive the reported reading time effects in Hackl, Koster-Hale & Varvoutis (2012). *Journal of Semantics*, (3), 1–8.

Grice, H. P. (1989). *Studies in the Way of Words.* Harvard: Harvard University Press.

Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics*, volume 3. New York: Harvard University Press.

Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). "Some", and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, *116*(1), 42–55.

Hackl, M. (2000). *Comparative quantifiers.* PhD thesis, Massachusetts Institute of Technology.

Hackl, M., Koster-Hale, J., & Varvoutis, J. (2012). Quantification and ACD: Evidence from real-time sentence processing. *Journal of Semantics*, *29*(2), 145–206.

Hartshorne, J., Liem Azar, S., Snedeker, J., & Kim, A. (2014). The neural computation of scalar implicature. *Language Cognition and Neuroscience*, *30*(620–634).

Haspelmath, M. (1997). *Indefinite Pronouns.* Oxford: Clarendon Press.

Heim, I. (2000). Degree operators and scope. In Jackson, B. & Matthews, T. (Eds.), *Semantics and Linguistic Theory*, volume 10.

Horn, L. (1972). *On the semantics and pragmatics of logical operators in English.* PhD thesis, UCLA, California.

Huang, Y. & Snedeker, J. (2009). Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive Psychology*, *58*(3), 376–415.

Jacobson, P. & Gibson, E. (2014). Processing of ACD gives no evidence for QR. In Young, R. B. (Ed.), *Semantics and Linguistic Theory (SALT)*, volume 24, (pp. 156–176).

Jansen, C. & Pollmann, M. (2001). On round numbers: Pragmatic aspects of numerical expressions. *Journal of Quantitative Linguistics*, *8*(3), 187–201.

Just, M. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354.

Kamp, H. (1973). Free choice permission. *Proceedings of the Aristotelian Society*, *74*, 57–74.

Katsos, N. & Bishop, D. V. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, *120*(1), 67–81.

Katsos, N. & Cummins, C. (2010). Pragmatics: From theory to experiment and back again. *Language and Linguistics Compass*, *4*(5), 282–295.

Kennedy, C. (2015). A "de-Fregean" semantics (and neo-Gricean pragmatics) for modified and unmodified numerals. *Semantics and Pragmatics*, *8*(10), 1–44.

Kratzer, A. & Shimoyama, J. (2002). Indeterminate pronouns: The view from japanese. In Otsu, Y. (Ed.), *The Proceedings of the Third Tokyo Conference on Psycholinguistics*, (pp. 1–25).

Krifka, M. (1999). At least some determiners aren't determiners. In K. Turner (Ed.), *The semantics/pragmatics interface from different points of view*, volume 1 (pp. 257–291). Oxford: Elsevier.

Krifka, M. (2009). Approximate interpretations of number words: A case for strategic communication. In E. Hinrichs & J. Nerbonne (Eds.), *Theory and Evidence in Semantics* (pp. 109–132). Stanford: CSLI Publications.

Kurtzman, H. S. & MacDonald, M. C. (1993). Resolution of quantifier scope ambiguities. *Cognition*, *48*(3), 243–279.

Lauer, S. (2009). Free relatives with -ever: Meaning and use. Unpublished Manuscript, Department of Linguistics, Stanford University, California, USA.

Lauer, S. (2012). Some news about *irgend* and *algún*. Ms. Stanford University.

Lauer, S. (2013). *Towards a dynamic pragmatics*. PhD thesis, Stanford University, Stanford, CA.

Lauer, S. (2014). Mandatory implicatures in Gricean pragmatics. In Degen, J., Franke, M., & Goodman, N. D. (Eds.), *Proceedings of the Formal & Experimental Pragmatics Workshop*, (pp. 21–28).

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

Lewis, S. (2013). *Pragmatic enrichment in language processing and development*. PhD thesis, University of Maryland, College Park, Maryland.

MacDonald, M. C., Just, M., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24*(1), 56–98.

Mayr, C. (2013). Implicatures of modified numerals. In I. Caponigro & C. Cecchetto (Eds.), *From grammar to meaning: The spontaneous logicality of language* (pp. 139–171). Cambridge: Cambridge University Press.

Mayr, C. & Meyer, M.-C. (2014). More than at least. *Two Days At Least* Workshop, Utrecht University.

McNabb, Y. & Penka, D. (2014). The interpretation of superlative modifiers and deontic modals: An experimental investigation. In Etxeberria, U., Fălău, A., Irurtzun, A., & Leferman, B. (Eds.), *Sinn und Bedeutung*, volume 18, (pp. 271–288).

McNabb, Y. & Penka, D. (2015). An experimental investigation of ignorance inferences and authoritative interpretations of superlative modifiers. Under review.

Mendia, J. A. (2016a). Focusing on scales: Superlatives as scalar modifiers. In Hammerly, C. & Prickett, B. (Eds.), *North Eastern Linguistics Society (NELS)*, volume 46, (pp. 11–24). Amherst, MA: GLSA Publications.

Mendia, J. A. (2016b). Known unknowns: epistemic inferences of superlative modifiers. Unpublished Manuscript, University of Massachusetts Amherst, Massachusetts, USA.

Mendia, J. A. (2016c). Reasoning with partial orders: Restrictions on ignorance inferences of superlative modifiers. In Moroney, M., Little, C.-R., Collard, J., & Burgdorf, D. (Eds.), *Semantics and Linguistic Theory (SALT)*, volume 26, (pp. 489–505).

Meyer, M.-C. (2013). *Ignorance and Grammar*. PhD thesis, MIT, Massachusetts.

Murray, W. S. & Rowan, M. (1998). Early, mandatory, pragmatic processing. *Journal of Psycholinguistic Research*, *27*(1), 1–22.

Myers, J. L., O'Brien, E. J., Albrecht, J. E., & Mason, R. A. (1994). Maintaining global coherence during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 876–886.

Nickel, B. (2011). Generically free choice. Unpublished Manuscript, Harvard University, Boston, USA. Manuscript, Harvard.

Nieuwland, M. S., Ditman, T., & Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language*, *63*(3), 324–346.

Nouwen, R. (2010). Two kinds of modified numerals. *Semantics and Pragmatics*, *3*(3), 1–41.

Nouwen, R. (2015). Modified numerals: The epistemic effect. In L. Alonso-Ovalle & P. Benito-Ménendez (Eds.), *Epistemic indefinites: Exploring modality beyond the verbal domain* chapter 11, (pp. 244–266). Oxford: Oxford University Press.

Nouwen, R., Alexandropoulou, S., & McNabb, Y. (2018). Experimental work on the semantics and pragmatics of modified numerals. In C. Cummins & N. Katsos (Eds.), *Handbook of experimental semantics and pragmatics*. Oxford: Oxford University Press.

Noveck, I. & Reboul, A. (2008). Experimental pragmatics: a Gricean turn in the study of language. *Trends in Cognitive Sciences*, *12*(11), 425–431.

Noveck, I. A. & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language*, *85*(2), 203–210.

Panizza, D., Chierchia, G., & Clifton, C. (2009). On the role of entailment patterns and scalar implicatures in the processing of numerals. *Journal of memory and language*, *61*(4), 503–518.

Pantazi, M., Kissine, M., & Klein, O. (2018). The power of the truth bias: False information affects memory and judgment even in the absence of distraction. *Social Cognition*, *36*(2).

Politzer-Ahles, S. & Fiorentino, R. (2013). The realization of scalar inferences: Context sensitivity without processing cost. *PloS ONE*, *8*(5).

Politzer-Ahles, S., Fiorentino, R., Jiang, X., & Zhou, X. . (2012). Distinct neural correlates for pragmatic and semantic meaning processing: An event-related potential investigation of scalar implicature processing using picture-sentence verification. *Brain Research*, *1490*, 134–152.

Politzer-Ahles, S. & Gwilliams, L. (2015). Involvement of prefrontal cortex in scalar implicatures: evidence from magnetoencephalography. *Language, Cognition and Neuroscience*, *30*(7), 853–866.

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory and Cognition*, *5*(5), 443–448.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy*, *27*, 367–391.

Sauerland, U. & Schumacher, P. B. (2016). Pragmatics: Theory and experiment growing together. *Linguistische Berichte*, *245*, 3–24.

Schwarz, B. (2012). *At least*: alternatives and symmetry. Paper presented at Workshop on Quantity Expressions, Universität Konstanz.

Schwarz, B. (2016a). Consistency preservation in quantity implicature: The case of at least. *Semantics and Pragmatics*, *9*(1), 1–47.

Schwarz, B. (2016b). *At least* and ignorance: A reply to Coppock & Brochhagen 2013. *Semantics & Pragmatics*, *9*(10), 1–17.

Schwarz, B. & Shimoyama, J. (2011). Negative islands and obviation by *wa* in Japanese degree questions. In Li, N. & Lutz, D. (Eds.), *Semantics and Linguistic Theory (SALT)*, volume 20, (pp. 702–719).

Simons, M. (2005). Dividing things up: The semantics of *or* and modal/*or* interaction. *Natural Language Semantics*, *13*(3), 271–316.

Spector, B. (2014). Global positive polarity items and obligatory exhaustivity. *Semantics and Pragmatics*, *7*(11), 1–61.

Spector, B. (2015). Why are class b modifiers global PPIs? Hurford disjunctions as a model for class b modifiers. Handout at Workshop on Negation and Polarity in Jerusalem.

Sprouse, J. (2011). A validation of Amazon Mechanical Turk for the collection of acceptability judgements in linguistic theory. *Behavioral Research Methods*, *43*, 155–167.

Sprouse, J., Schütze, C., & Almeida, D. (2013). A comparsion of informal and formal acceptability judgments using a random sample from *Linguistic Inquiry* 2001–2010. *Lingua*, *134*, 219–248.

Spychalska, M. (2015). At least not false, at most possible: between truth and assertibility of superlative quantifiers. *Synthese*, *195*(2), 1–32.

Spychalska, M., Kontinen, J., & Werning, M. (2016). Investigating scalar implicatures in a truth-value judgement task: evidence from event-related brain potentials. *Language, Cognition and Neuroscience*, *31*(6), 817–840.

Starr, W. B. (2016). Expressing permission. In Moroney, M., Little, C.-R., Collard, J., & Burgdorf, D. (Eds.), *Semantics and Linguistic Theory (SALT)*, volume 26, (pp. 325–349).

Szabolcsi, A. (2014). Quantification and ACD: What is the evidence from real-time processing evidence for? A response to Hackl et al. (2012). *Journal of Semantics*, *31*(1), 135–145.

Tomlinson, J. M., Bailey, T. M., & Bott, L. (2013). Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of Memory and Language*, *69*(1), 18–35.

Tunstall, S. (1998). *The interpretation of quantifiers: Semantics and processing.* PhD thesis, University of Massachusetts Amherst, Massachusetts.

van Berkum, J. J., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, *11*(6), 657–671.

van Berkum, J. J., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research*, *17*(3), 701–718.

van Tiel, B., van Miltenburg, E., Zevakhina, N., & Geurts, B. (2016). Scalar diversity. *Journal of Semantics*, *33*(1), 137–175.

von Fintel, K. (2000). Whatever. In Jackson, B. & Matthews, T. (Eds.), *Semantics and Linguistic Theory (SALT)*, volume 10, (pp. 27–40).

Westera, M. & Brasoveanu, A. (2014). Ignorance in context: The interaction of modified numerals and QUDs. In Snider, T., D'Antonio, S., & Weigand, M. (Eds.), *Semantics and Linguistic Theory (SALT)*, volume 24, (pp. 414–431).

Zimmermann, T. E. (2000). Free choice disjunction and epistemic possibility. *Natural Language Semantics*, *8*(4), 255–290.

Samenvatting in het Nederlands

De pragmatiek van telwoord-modificeerders

Dit proefschrift omvat een studie van de pragmatiek van telwoord-modificeerders. Nouwen (2010) onderscheidt twee klassen modificeerders van telwoorden. De expressies in de twee klassen zijn waarheids-conditioneel gezien equivalent, maar verschillen doordat klasse A modificeerders geen onwetendheid-effecten laten zien, terwijl expressies in klasse B dat wel doen. Binnen het gros van de relevante literatuur wordt dit onderscheid verklaard in termen van pragmatische mechanismen. Recentelijk is het onderscheid echter in twijfel getrokken door een klein aantal studies. Ciardelli, Coppock & Roelofsen (2017) en Westera & Brasoveanu (2014) (zie ook ongepubliceerd werk van Mayr & Meyer, 2014) beargumenteren dat klasse A modificeerders eveneens onwetendheid-effecten laten zien. Naast onwetendheid-gevolgtrekkingen, zijn er verdere effecten die met het gebruik van telwoord-modificeerders samenhangen, met name variatie-effecten en gevolgtrekkingen van onverschilligheid. Deze effecten hebben nauwelijks aandacht gekregen in de literatuur.

Deze dissertatie onderzoekt vanuit deze achtergrond alle drie de typen gevolgtrekkingen die we bij gemodificeerde telwoorden zien, voor modificeerders van zowel de klasse A als de klasse B variant. Het doel is om er achter te komen waar de grens ligt tussen semantische en pragmatische aspecten van betekenis voor elk van de twee klassen, en in hoeverre de twee klassen daadwerkelijk – in zowel semantisch als pragmatisch opzicht – verschillen.

De onderzoeksvragen afgeleid van deze doelstelling worden in hoofdstuk 1 geïntroduceerd:

- ▷ Met welke waarschijnlijkheid vinden variatie-gevolgtrekkingen plaats? Zijn er tussen verschillende modificeerders verschillen in de mate waarin variatie-effecten voorkomen? Wat is de semantische/pragmatische status en kracht van variatie-effecten? Welke precieze kenmerken hebben dit soort gevolgtrekkingen en wat betekent dat voor de mechanisme dat verantwoordelijk

is voor het afleiden van die inferenties?

- ▷ Hoe vinden gevolgtrekkingen van onwetendheid plaats tijdens de incrementale interpretatie van gemodificeerde telwoorden? Wat kunnen we aan de hand van zo'n verwerkings-profiel concluderen over de aard en kracht van deze gevolgtrekkingen? En wat leert dat profiel ons over het verantwoordelijke mechanisme? Wat is de relatie tussen dat mechanisme en het mechanisme verantwoordelijk voor variatie-gevolgtrekkingen?

- ▷ Is spreker-onverschilligheid een beschikbare implicatie van gemodificeerde telwoorden? Wat is de relatie tussen onverschilligheid en onwetendheid? Wat is het verwerkingsprofiel van onverschilligheid en hoe verhoudt dat zich ten opzichte van onwetendheid?

Hoofdstuk 2 bevat een overzicht van de flinke theoretische literatuur op het gebied van onwetendheid-effecten bij gemodificeerde telwoorden, maar ook van de kleine hoeveelheid werk over variatie-gevolgtrekkingen. Centraal in die literatuur staat de oppositie tussen superlatieve en comparatieve modificeerders (respectievelijk klasse B en klasse A). Het doel van dit hoofdstuk is het identificeren van de theoretische aspecten die cruciaal zijn voor het experimentele werk in de rest van de dissertatie.

Hoofdstuk 3 richt zich op de eerste set onderzoeksvragen (zie hierboven), betreffende variatie-effecten. Het bevat nieuwe data verkregen via een serie off-line experimenten (afgenomen in het Nederlands) naar zowel klasse B expressies ("minstens", "n of meer", "maximaal", "n of minder"), als klasse A expressies ("meer dan", "minder dan"). In die experimenten onderzoek ik de waarschijnlijkheid en kracht van variatie-effecten bij telwoord-modificeerders in het bereik van een universele kwantor. Ook de specifieke aard van variatie-gevolgtrekkingen (ondergespecificeerde variatie versus specifieke partiële variatie) wordt met deze experimenten bepaald. Alles samengenomen tonen de experimenten aan dat superlatieve, disjunctieve en ook comparatieve telwoord-modificeerders optioneel tot variatie-implicaturen leiden. Het gaat daarbij om specifieke partiële variatie, gekarakteriseerd door een inferentie over de minimale waarde n die compatibel is met het gemodificeerde telwoord. Hoewel zowel superlatieve als comparatieve modificeerders deze implicaturen teweegbrengen, blijken de implicaturen bij comparatieve modificeerders van een zwakkere aard te zijn. De experimentele resultaten hebben consequenties voor het pragmatische mechanisme verantwoordelijk voor de implicaturen. In het bijzonder tonen de experimenten aan dat dit mechanisme zal moeten leunen op twee scalaire alternatieven, namelijk $[n]$ (de precieze waarde van het telwoord) en $[n+1, \ldots)$ (het deel van de schaal dat start bij het telwoord plus één).

Hoofdstuk 4 geeft een gedetailleerd overzicht van de bestaande literatuur op het gebied van gevolgtrekkingen van spreker-onwetendheid met betrekking tot superlatieve en comparatieve modificeerders. Dit vormt de basis voor het experimentele werk dat volgt in hoofdstuk 5 en 6. In die bespreking komt naar voren dat de bevindingen van het merendeel van de studies laten zien dat

gevolgtrekkingen van onwetendheid alleen beschikbaar zijn in combinatie met superlatieve modificeerders en dat deze een pragmatische, in plaats van een semantische, status hebben. De studie van Westera & Brasoveanu (2014) krijgt hier speciale aandacht omdat deze als eerste beargumenteerde en liet zien dat comparatieve modificeerders onder bepaalde omstandigheden ook onwetendheidseffecten kunnen doen ontstaan. Ook wordt besproken dat het bewijs tot op heden veelal van indirecte aard is en verder dat het weinige directe bewijs dat er wel is, controversieel te noemen is. Op basis van deze uiteenzetting wordt geconcludeerd dat een meer nauwkeurige en directe studie van gevolgtrekkingen van onwetendheid noodzakelijk is. Hierbij moeten gekeken worden naar de incrementele interpretatie van telwoord-modificeerders. Een dergelijke studie zou licht kunnen werpen op de onderliggende processen die verantwoordelijk zijn voor het afleiden van gevolgtrekkingen van spreker-onwetendheid bij zowel superlatieve als comparatieve modificeerders, en zou in het bijzonder ook helpen bij het beoordelen van de verscheidene pragmatische voorstellen over onwetendheid.

In hoofstuk 5 en 6 staan gevolgtrekkingen van spreker-onwetendheid centraal en in het bijzonder hoe deze gevolgtrekkingen plaatsvinden tijdens incrementele interpretatie. Het doel van deze hoofdstukken is antwoord te geven op de tweede set onderzoeksvragen hierboven. Hoofdstuk 5 rapporteert over een eyetrackingexperiment dat gebruik maakt van het lees-paradigma en in het Nederlands werd uitgevoerd. Dit experiment onderzoekt niet-ingebedde en ingebedde voorbeelden van de superlatieve modificeerder *minstens* in een context of met een spreker die beperkte kennis heeft (onwetend) of met een spreker met volledige kennis (autoriteit). De manier waarop de contexten opgebouwd zijn is in beginsel geïnspireerd door Breheny, Katsos & Williams (2006). Deze studie, en een groot aantal daaropvolgende studies, demonstreerde de contextgevoeligheid van de berekening van scalaire implicaturen, in contrast met een standaard benadering van hoe implicaturen worden berekend. Gezien deze robuuste bevindingen is het idee achter onze contextmanipulatie de volgende: De context met de onwetende spreker zou de berekening van een specifieke onwetendheid-implicatuur in een targetzin met *minstens* moeten bevorderen, terwijl een dergelijke implicatuur minder waarschijnlijk zou zijn in een context met een spreker als autoriteit. Interpretaties van spreker-onwetendheid van ingebedde gevallen van superlatieve modificeerders ontstaan volgens de meerderheid van de theoretische verhandelingen volgens een *wide-scope* beweging van de superlatieve modificeerder in verhouding tot de huidige operator. Daarom hebben we eveneens de superlatieve modificeerder in ingebedde modale contexten getest, opdat we konden kijken of we met eyetracking sporen van *wide scope* bewegingen van *minstens* konden vinden. Aangezien de verschillende verklaringen van spreker-onwetendheid overwegend van theoretische aard zijn zonder psycholinguistische claims, namen wij een neutraal standpunt in ten opzichte van de verwerking van de superlatieve modificeerder en de online totstandkoming van zijn specifieke spreker-onwetendheidsinferenties en overwogen verschillende verwerkingsmogelijkheden. We vonden een verwerkingsboete in de

waarschijnlijkheid van herlezen in de regio die de telwoordfrase bevatte (e.g., "zes mensen") volgend op de telwoord-modificeerder regio in de onwetendheidscontext. We vonden geen interactie tussen de Context en Werkwoord factoren. Dat wil zeggen, we vonden geen bewijs dat onwetendheid met *minstens* in combinatie met een universele modal een extra operatie inhoudt, zoals bijvoorbeeld *covert movement*.

Samengenomen hebben we onze bevindingen geïnterpreteerd als suggererend dat effecten van spreker-onwetendheid van superlatieve modificeerders niet-verplichte, context-afhankelijke implicaties zijn van het specifieke type onwetendheid en afgeleid worden via een pragmatisch mechanisme zoals dat gespecificeerd door de op kwantiteit gebaseerde Griceaanse verklaringen van onwetendheid. Ik concludeer dat de derivatie van spreker-onwetendheidsinterpretaties gepaard gaat met een verwerkingsboete, zoals te zien in de late maten. Met dit experiment heb ik het eerst deel van de set vragen met betrekking tot onwetendheidsinferenties (zie hierboven) beantwoord. Een aantal alternatieve verklaringen van het gevonden effect dat we toeschreven aan de derivatie van spreker-onwetendheid blijft echter nog over. Daarom is het noodzakelijk een vervolgexperiment uit te voeren.

Hoofdstuk 6 presenteert dit vervolgexperiment. In dit experiment worden een aantal resterende uitdagingen opgelost en proberen we meer te weten te komen over of het effect geobserveerd in het vorige eyetrackingexperiment inderdaad toe te schrijven was aan de derivatie van een specifieke onwetendheid-implicatuur. Meer concreet, met behulp van feedback door moedertaalsprekers verken ik de mogelijkheid dat de informatie van de targetzin met *minstens* te specifiek was, en daarom *infelicitous*, gegeven de voorgaande verklaring van onwetendheid door de onwetende spreker en dat dit de reden was voor de verwerkingsboete. Ik beargumenteer dat wat er specifiek is aan de targetzin met de superlatieve modificeerder bestond uit (i) de kernbetekenis van *minstens n* als het uitsluiten van waardes lager dan n, of (ii) de specifieke onwetendheid-implicatuur geassocieerd met *minstens*. Om uit te vinden of het geobserveerde Contexteffect toegeschreven kan worden aan de derivatie van de specifieke onwetendheid-implicatuur – online ofwel offline door de verstoring vanwege de *infelicity* op basis van (ii) – of dat het Contexteffect resulteerde door de verstoring vanwege de *infelicity* door (i), heb ik eveneens de comparatieve modificeerder *meer dan* getest. De comparatieve conditie zou helpen deze verschillende mogelijkheden te ontwarren omdat *meer dan* een even specifieke kernbetekenis heeft als *minstens* en tegelijkertijd aangenomen wordt dat het geassocieerd is met een minder robuuste, specifieke onwetendheid-implicatuur dan *minstens*. Daarom werd verwacht dat als het Contexteffect zoals geobserveerd in de eerdere experimenten toe te schrijven zou zijn aan de *infelicity* veroorzaakt door de specifieke kernbetekenis van het gemodificeerde telwoord op basis van de eerdere onwetendheidscontext, beide telwoord-modificatie condities zich hetzelfde zouden moeten gedragen. Echter, de twee telwoord-modificeerders zouden een verschillend verwerkingsprofiel moeten laten zien als het geobserveerde Contexteffect toe te schrijven zou zijn aan de (online of

offline) derivatie van de specifieke onwetendheid-implicatuur.

We repliceerden het Contexteffect uit het eerste eyetracking-experiment: Lezers waren geneigd de regio van de telwoordfrase (*tien mensen*) vaker te herlezen binnen een onwetendheidcontext dan binnen een autoriteitscontext. Cruciaal was dat zij dit vaker deden in de superlatieve conditie dan in de comparatieve conditie. Over het algemeen was het verwerkingsprofiel van de comparatieve conditie anders dan dat van de superlatieve conditie. Een verwerkingsboete die ten koste ging van de comparatieve conditie van onwetendheiditems werd geobserveerd bij de herleesduur van de regio volgend op de gemodificeerde telwoordfrase (bijv. *met veel oog voor detail*). Aan de hand van deze bevindingen concludeer ik in hoofdstuk 6 dat het Contexteffect – dat wordt aangetoond in het eerste eyetracking-experiment en gerepliceerd in het tweede experiment – toegewezen zou moeten worden aan de derivatie van een optionele specifieke spreker-onwetendheid-implicatuur van de telwoordfrase die gemodificeerd wordt door *minstens* in een spreker-onwetendheid-context – ofwel online ofwel indirect gemanifesteerd door het niet passen van de resulterende interpretatie van de targetzin bij de voorgaande onwetendheiduiting.

De verwerkingsboete die werd geobserveerd bij de prepositiefrase na het gemodificeerde telwoord zie ik als een indicatie van een ander pragmatisch proces dat zich op een later punt in de targetzin manifesteert en dat ten koste gaat van de comparatieve items (in de onwetendheidconditie). Ofwel wordt er een zwakkere specifieke onwetendheid-implicatuur en/of van een ander type gegenereerd bij de comparatieve modificeerder vergeleken met de superlatieve modificeerder (online of indirect gemanifesteerd, met dezelfde reden als bij de superlatieve modificeerder), of lezers voltrokken bij nader inzien een soort *Manner*-achtige redenering nadat zij een comparatieve modificeerder tegenkwamen in een onwetendheidcontext: Waarom gebruikte de spreker/schrijver niet *minstens*, wat een betere indicatie van onwetendheid zou zijn? Dit resulteert in de derivatie van een *Manner*-implicatuur, die online werd aangetoond. Uiteindelijk beweer ik dat, gegeven de verklaring gebaseerd op *Manner*, *minstens* een betere aanwijzing is van onwetendheid omdat het geassocieerd wordt met een sterkere onwetendheid-implicatuur.

Verder stel ik voor om het aangetoonde sterkteverschil in specifieke onwetendheid-inferenties tussen de twee types telwoord-modificeerders en tussen hun niet-verplichte en contextafhankelijke status te ondervangen door de zogenaamde *Composite Theory*. Deze theorie omvat Nouwens (2015) verklaring voor superlatieve modificeerders en een standaard verklaring van comparitieve modificeerders gebaseerd op *Quantity*. Dat wil zeggen, superlatieve en comparatieve modificeerders worden geassocieerd met dezelfde verzameling alternatieven, zoals hieronder geïllustreerd:

(1)

$$\left\langle \begin{array}{lllll} \ldots & [n+2] & [n+1] & [n] & \\ & & & & [\text{n}, \ldots) \\ \ldots & [\text{n}{+}3, \ldots) & [\text{n}{+}2, \ldots) & [\text{n}{+}1, \ldots) & \end{array} \right\rangle$$

(n is de minimale waarde die compatibel is met het gemodificeerde telwoord, $[n]$ staat voor de propositie met *only/exactly* n en $[n,)$ staat voor de propositie met het gemodificeerde telwoord). Belangrijk is dat ondanks dat de twee types telwoord-modificeerders dezelfde alternatieven hebben, dat deze deels verschillend zijn vanwege hun bron. De *exactly* alternatieven zijn op basis van presupposities voor superlatieve modificeerders en zijn input voor een mechanisme gebaseerd op *Manner*, terwijl die alternatieven bij comparatieve modificeerders tot stand komen via een *Horn*-schaal en gebruikt worden in een standaard redenering gebaseerd op *Quantity*. De overige alternatieven komen tot stand via de *Horn*-telwoordschaal en zijn onderhevig aan een proces gebaseerd op *Quantity* voor zowel superlatieve als comparatieve modificeerders. Afhankelijk van of de telwoord-modificeerders voorkomen in ingebedde omgevingen (bijv. omgevingen met universele kwantoren) of niet-ingebedde omgevingen, is de output van het relevante pragmatische mechanisme specifieke gedeeltelijke variatie en specifieke spreker onwetendheid-implicaturen, respectievelijk. Met dit voorstel sloten we ons onderzoek naar spreker-onwetendheid-effecten van gemodificeerde telwoorden af, waarmee we ook een adequaat antwoord gaven op de tweede set onderzoeksvragen, die te maken had met onwetendheid (zie boven).

Tot slot heeft hoofdstuk 7 als doel om de laatste verzameing vragen, die te maken had met spreker-onverschilligheid, te beantwoorden. Aan de hand van een aantal taalkundige toetsen en een offline onderzoek, beargumenteer ik dat superlatieve en comparatieve telwoord-modificeerders ook kunnen leiden tot spreker-onverschilligheid-effecten, zodanig dat de spreker overweegt dat het niet relevant is om het exacte aantal onder overweging te specificeren. In een lijn met onderzoeken naar spreker-onverschilligheid uit andere empirische domeinen, stel ik voor spreker-onverschilligheid te verklaren op eenzelfde manier als spreker-onwetendheid, door de *Composite Theory* uit te breiden zodanig dat het beide types inferenties op een uniforme wijze verklaart als (primair) *Quantity*-implicaties. Eigenlijk behoud ik het mechanisme voor elk type modificeerder zoals hierboven beschreven, maar vervang de geloof-operator, $\Box_{Bel}$, die verantwoordelijk is voor de derivatie van onwetendheidimplicaties, door een algemenere operator, te weten $\Box_{Assert}$. $\Box_{Assert}p$ staat voor *de spreker gelooft dat p waar is en vindt p relevant.* Vandaar zijn gegeven een uiting $\Box_{Assert}[n, ...)$ de resulterende onwetendheid- en spreker-onverschilligheid-implicaties zodanig dat de sterkere proposities $[n]$ en $[n+1,)$ niet beweerbaar omdat de spreker niet gelooft dat ze waar zijn en ze niet relevant vindt. Net als de *Composite Theory* voorspelt de uitgebreide versie van de *Composite Theory* dat deze implicaties robuster aanwezig zijn voor superlatieve dan voor comparatieve modificeerders.

De voorspellingen van de *extended Composite Theory* werden met hulp van een extra context conditie (onverschilligheidsconditie) getest in ons tweede eyetracking-experiment. De resultaten van de onverschilligheidsconditie verschillen echter van de resultaten van de onwetendheidsconditie, wat een probleem is voor de *extended Composite Theory*. Er was geen verschil tussen de twee

telwoord-modificeerdes in de regio's waar de relevante effecten in de onwetendheidsconditie geobserveerd werden. Sterker nog, er was een indicatie dat onverschilligheid sterker is bij comparatieve modificeerders (telwoord-modificeerder regio). De *extended Composite Theory* kan alleen in stand worden gehouden als de onwetendheidsconventie van superlatieve modificeerders ingebouwd wordt, omdat comparatieve modificeerders betere signalen voor onverschilligheid blijken te zijn dan superlatieve modificeerders. Beide telwoord-modificeerders kunnen onwetendheids- en onverschilligheid-effecten hebben, zoals voorspeld door de *extended Composite Theory*, met het verschil dat onwetendheid-effecten van superlatieve modificeerders in het process zijn een conventie te worden. Dit kan als gevolg hebben dat comparatieve modificeerders de functie overnemen en dus betere signalen blijken voor onverschilligheid (in een niet-ingebedde context). Het kan echter ook zijn dat de effecten in de onverschilligheidsconditie komen door scalaire implicaturen die te maken hebben met niveau's van fijnmazigheid, een ander soort gevolgtrekking die in een niet-ingebedde context voorkomt met een co-operative spreker met de nodige kennis (Cummins, Sauerland & Solt, 2012). Hoe de resultaten van de onverschilligheidsconditie in het tweede eyetracking-experiment ook geinterpreteerd worden, ik beweer dat de *extended Composite Theory* in stand kan worden gehouden (als conventie ingebouwd wordt). Met dit hoofdstuk hebben we dus een compleet beeld van de pragmatiek van telwoord-modificeerders.

De contributie van deze proefschrift

Op theoretisch niveau

Door naar drie verschillende effecten van telwoord-modificeerders te kijken – variatie, onwetendheid en onverschilligheid – draagt dit proefschrift bij aan een compleet en globaal beeld van de pragmatiek van telwoord-modificeerders. Een hoofd-resultaat is dat superlatieve telwoord-modificeerders context-afhankelijke, niet verplichte gevolgtrekkingen van onwetendheid genereren. Dit helpt bij het evalueren van de veelvoudige verklaringen van onwetendheid-effecten van superlatieve modificeerders in de literatuur. Onze resultaten zijn in lijn met benaderingen die gebouwd zijn op kwantiteits-overwegingen, zoals Büring (2008), Cummins & Katsos (2010), Kennedy (2015), Nouwen (2015) en Schwarz (2016a), maar niet met Ciardelli et al. (2017), Cohen & Krifka (2014), Coppock & Brochhagen (2013b), Geurts & Nouwen (2007), Nouwen (2010), Spector (2015) of Spychalska (2015).

Een ander hoofd-resultaat is dat de twee klassen van telwoord-modificeerders, superlatieve en comparatieve modificeerders, representatief voor de klassen A en B van modificeerders, dezelfde pragmatieke gevolgtrekkingen genereren. Dit is in lijn met onder ander de voorstellen van Ciardelli et al. (2017), Mayr & Meyer (2014) en Westra & Brasoveanu (2014) over onwetendheid-effecten en met Mayr (2013) over variatie-effecten, maar niet met Büring (2008), Coppock

& Brochhagen (2013b), Cummins & Katsos (2010), Geurts & Nouwen (2007), Kennedy (2015), Nouwen (2010) en Spychalska (2015) over onwetendheid-effecten en Büring (2008), Coppock & Brochhagen (2013b) en Nouwen (2015) over variatie-effecten. We hebben verder aangetoond dat superlatieve en comparatieve modificeerders, ook genereren ze dezelfde gevolgtrekkingen, dit in verschillende mate doen. Dit suggereerd dat Nouwen's (2010) classificatie van klasse A en B modificeerders (en de herformulering hiervan in Nouwen (2015)) geen classificatie is van telwoord-modificeerders met betrekking tot de aan- of afwezigheid van gevolgtrekkingen, maar een classificatie gebasseerd op de mate of waarschijnlijkheid van het voorkomen van de gevolgtrekkingen.

Deze observatie van verschillen in mate van het genereren van gevolgtrekkingen doet denken aan de resultaten van Doran, Baker, McNabb, Larson & Ward (2009), Doran, Ward, Larson, McNabb & Baker (2012) en van Tiel, van Miltenburg, Zevakhina & Geurts (2016) over de verschillende mate waarmee scalaire termen gevolgtrekkingen genereren. Deze observatie is meestal niet meegenomen in formele theorieën. De *extended Composite Theory* zoals voorgesteld in dit proefschrift kan niet alleen de pragmatieke betekenis van de gevolgtrekkingen verklaren, bovenop de semantieke betekenis, maar ook het verschil in de mate waarin superlatieve en comparatieve modificeerders deze gevolgtrekkingen genereren. Ciardelli et al. (2017) geven een soortgelijke verklaring, maar dan alleen voor het verschill tussen klasse A en klasse B modificeerders met betrekking tot onverschilligheid-gevolgtrekkingen. De context-afhankelijke gevolgtrekkingen van onwetendheid die Ciardelli et al. voor superlatieve modificeerders voorstellen is niet in lijn met de resultaten van dit proefschrift.

Terug naar de vraag van het begin: Waarom heeft taal verschillende manieren om dezelfde betekenis uit te drukken? Of, om het iets specifieker te zeggen, waarom heeft taal één uitdrukking die altijd gevolgtrekkingen genereerd (superlatieven) en één uitdrukking die dezelfde gevolgrekkingen genereerd, maar dan in mindere mate (comparatieve)? Het is hier dat mijn voorstel van de *extended Composite Theory* met de ingebouwde onwetendheidsconventie van superlatieve modificeerders belangrijk wordt. Dit voorstel blijkt namelijk een antwoord te geven op deze vraag met betrekking tot zogenaamde *division of pragmatic labor* tussen superlatieve en comparatieve telwoord-modificeerders en de gevolgtrekkingen die deze in niet-ingebedde contexts genereren: Door het conventie-process van onwetendheid-effecten met superlatieve modificeerders, worden deze onwetendheid-effecten met superlatieve modificeerders geassocieerd, en daarom zijn comparatieve modificeerders betere signalen voor onverschilligheid-effecten. Dit zou door toekomstig onderzoek nog duidelijker moeten worden.

Een aantrekkelijk idee voor verder onderzoek zou zijn om een diachroon corpusonderzoek naar telwoord-modificeerders uit te voeren. Een dergelijk onderzoek maakt het mogelijk om de verschillende toepassingen en effecten van elke telwoord-modificeerder (variatie, onwetendheid en onverschilligheid) over opeenvolgende perioden te annoteren en monitoren, zodat de bovengenoemde suggestie getoetst kan worden. Omdat de robuustheid van onwetendheid-impli-

caturen een cross-linguïstisch fenomeen is (Nouwen, 2010), zou het interessant zijn om dat corpusonderzoek in verschillende talen uit te voeren. Een dergelijk grootschalig diachroon onderzoek is uitgevoerd door Aloni, Aguilar-Guevara, Port, Simik, Solt, de Vos & Zeijlstra (zie, bijvoorbeeld, Aloni, 2016) om de verschillende toepassingen van free-choice indefinieten, zoals free-choice en spreker-onwetendheid-effecten in het Spaans, Nederlands en Duits te verkennen. Hun onderzoek toonde aan dat free-choice inferenties van het Spaanse indefiniete *cualquiera* en van het Nederlandse indefiniete *wie dan ook* worden afgeleid als semantische *entailments*, terwijl het versteende implicaturen zijn, als gevolg van semantische verandering, met het Duitse indefiniet *irgend/irgendein*. Wat onwetendheid-gevolgtrekkingen betreft kwamen ze tot de conclusie dat die het resultaat zijn van lexicaal gecodeerde condities op het gebruik van de expressies. Het moge duidelijk zijn dat diachroon corpusonderzoek significante inzichten kan geven in de huidige status van verschillende effecten die worden geassocieerd met bepaalde uitdrukkingen en, belangrijker, in hoe dit tot stand is gekomen in de loop van de tijd voor verschillende talen. Een dergelijke lijn van onderzoek kan daarom duidelijk maken of vergelijkbare conclusies als die van Aloni (2016) getrokken kunnen worden voor de variatie en onwetendheid-effecten van telwoord-modificeerders, en of de voorgestelde *extended Composite Theory* – waarin beide, evenals onverschilligheid-effecten, gevat worden – en mijn aanvullende diachrone suggestie op het juiste spoor zitten.

Gegeven de conclusies tot nu toe, zou het duidelijk moeten zijn dat telwoord-modificeerders een aantal verschillende inferenties oproepen, te weten variatie, spreker-onwetendheid en spreker-onverschilligheid, en dat we zouden moeten zoeken naar een uitgebreid theoretisch kader om rekening te houden met verschillen in waarschijnlijkheid/kracht van de inferenties, zoals die worden gezien tussen superlatieve en comparatieve modificeerders (met betrekking tot de voornoemde inferenties), waarbij we mogelijk ook rekening moeten houden met conventionalisatieprocessen.

Op psycholinguïstisch niveau

De oogbeweging lees-studies waar in deze thesis verslag van wordt gedaan tonen – direct of indirect – sporen van de derivatie van onwetendheid-gevolgtrekkingen, van onverschilligheids- of op granulariteit gebaseerde scalaire gevolgtrekkingen, en mogelijk ook van op het *Manner*-maxime gebasseere redeneringen. Deze bevindingen komen overeen met andere studies waarin kwantiteits- of informativiteitsgerelateerde fenomenen onderzocht werden (bijvoorbeeld Breheny et al., 2006; Engelhardt, Bailey & Ferreira, 2006; Fukumura & van Gompel, 2017; Huang & Snedeker, 2009; Panizza, Chierchia & Clifton, 2009), en waarin ook indicaties werden gevonden dat pragmatisch redeneren getuigt van incrementele interpretatie. Onze relevante bevindingen lijken bovendien te suggereren dat sterkere/meer robuuste inferenties vroeger in incrementele interpretatie voorkomen en dat ze, in het bijzonder, sterker verbonden zijn met de trigger, terwijl zwakkere inferenties in een regio veel later dan de trigger kun-

nen ontstaan, vgl. locatieverschillen van effecten die geassocieerd worden met de superlatieve vs. de comparatieve modificeerder.

Bovendien tonen onze online bevindingen aan dat pragmatisch redeneren zich pas laat in de oogbeweging manifesteert, dat wil zeggen, in late maten zoals de waarschijnlijkheid dat er herlezen wordt en in herleestijden, of in tussenliggende maten, zoals de duur van het regressiepad en totale leestijden. Dit ondersteunt het idee dat deze maten geassocieerd moeten worden met hoogniveau processen.

Tot slot gaan onze bevindingen in tegen een defaultistische visie op implicaturen, zoals die van Levinson (2000), waarvoor implicaturen standaard en moeiteloos berekend zouden moeten worden. Dit blijkt duidelijk uit de Context-effecten die we in de eyetracking-experimenten geobserveerd hebben, waarin te zien was dat er minder/geen implicaturen gegenereerd werden in de autoriteitscontextconditie, zoals ook ondersteund werd door de hogere acceptabiliteitsscores die de autoriteitsconditie kreeg in de relevante pretests. Eigenlijk zouden, volgens Levinson (2000), tegengestelde Context-effecten gevonden moeten worden in de eyetracking-experimenten, vanwege het wegvallen van de default onwetendheidsinterpretatie in de autoriteitscontext. Onze contextafhankelijke effecten sluiten daarentegen aan bij de Context-driven visie op implicaturen.

Al met al biedt onze studie bewijs uit een nieuw empirisch domein dat de berekening van pragmatische inferenties gevoelig is voor contextuele cues. Voorzover we de gevonden verwerkings-kosten kunnen zien als direct veroorzaakt door de online berekening van interpretaties van onwetendheid of onverschilligheid van de spreker en/of door een *Manner* implicatuur, wordt bovendien gesuggereerd dat dergelijke berekeningen moeite kosten. Dat ondersteunt de Context-driven visie op implicaturen, zoals bijvoorbeeld door Breheny et al. (2006) of Bott & Noveck (2004) gepresenteerd, waarbij de berekening van implicaturen moeite kost in de verwerking. De kloof tussen pragmatische theorie en taalgebruik is echter nog steeds aanzienlijk. Deze thesis heeft aangetoond dat de theorie fijnmaziger moet zijn om de data te kunnen verbinden.

Curriculum vitae

Stavroula Alexandropoulou was born on April 25, 1986 in Athens, Greece. She studied Greek Philology at the Aristotle University of Thessaloniki (AUTh) and graduated with a specialization in Linguistics in November 2009. In 2010, she moved to Utrecht to attend a research master's in Linguistics (UiL-OTS, Utrecht University). She graduated *cum laude* with a major in semantics/ pragmatics and a minor in experimental psycholinguistics in September 2013. She worked as a syntactic and semantic annotator between October 2011 and September 2013 in the NWO funded VICI project "Between Language and Common Sense", Utrecht University. In October 2013, she started her PhD at the UiL-OTS in Utrecht University in the ERC project "Restriction and Obviation in Scalar Expressions" (ROSE). This dissertation is the result of her PhD research.