Langage et logique : les cas des éléments à polarité négative et des implicatures scalaires

Milica Denić

PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'École Normale Supérieure

Langage et logique : les cas des éléments à polarité négative et des implicatures scalaires

Soutenue par

Milica DENIĆ

Le 21 juin 2019

Ecole doctorale n° 158

ED3C Cerveau Cognition Comportement

Spécialité

Sciences cognitives

Composition du jury :

David, BARNER
Full Professor, University of California, San Diego
Président, rapporteur

Gennaro, CHIERCHIA
Full Professor, Harvard University
Rapporteur

Angelika, KRATZER
Emeritus Professor, University of Massachusetts, Amherst
Examinatrice

Lyn, TIEU
Senior Research Fellow, Western Sydney University
Examinatrice

Benjamin, SPECTOR
Directeur de recherche, École Normale Supérieure
Directeur de thèse

Emmanuel, CHEMLA
Directeur de recherche, École Normale Supérieure
Co-Directeur de thèse

ENS

Université Paris Sciences et Lettres
École Normale Supérieure
Département d'études cognitives

Milica Denić

Language and logic: the cases of negative polarity items and scalar implicatures

Supervisors: Benjamin SPECTOR and Emmanuel CHEMLA

Ph.D. Thesis
Institut Jean-Nicod
Laboratoire de Sciences Cognitives et Psycholinguistique
École Doctorale 158 Cerveau Cognition Comportement
Defense date: 21st June 2019

Abstract

Numerous linguistic phenomena have been shown to correlate with some logical properties of the sentence with which they occur. Examples of such properties are logical entailments supported by the sentence, or the logical relation between it and some alternative sentence. In this thesis, we explore (A) whether these logical correlates play a causal role in the linguistic phenomena in question, and (B) at what level these computations of logical correlates are performed. There are two broad possibilities concerning (B): (i) these logical correlates could be computed in a formal system that does not have access to contextual knowledge, call this system grammar (Fox and Hackl, 2006, a.o.), or (ii) they could be computed post-grammatically, and therefore have access to contextual knowledge.

The first phenomenon we focus on is the licensing of negative polarity items (NPIs). NPIs are lexical items or phrases (e.g., 'lift a finger', 'anything') whose acceptability is sensitive to a logical property of their linguistic environment: namely, its "downward-monotonicity" (Fauconnier, 1975a; Ladusaw, 1979). An environment is downward-monotonic if it licenses inferences from sets (e.g., 'pie') to subsets (e.g., 'pumpkin pie'). Within this thesis, to advance on (A), we explore a phenomenon (the so-called intervention effects in NPI licensing) which is challenging to the proposal that the monotonicity of the environment plays a causal role in the acceptability of negative polarity items. Despite these challenges, an account has been developed which reduces intervention effects to monotonicity properties of the environment after all (Chierchia, 2004). We experimentally test predictions of this account. Our findings offer both some evidence in favor of and some challenges to the monotonicity approach to intervention effects, thus raising new questions about intervention effects in NPI licensing. Next, to advance on (B), we explore the question of whether NPIs influence inferential judgments of monotonicity, finding that they do. These results, together with the results by Chemla et al. 2011 and empirical arguments put forward in Crnič 2014, suggest that the licensing of NPIs is indeed sustained by post-grammatical computations of monotonicity.

The second phenomenon that we focus on is scalar implicatures. In certain cases, when two sentences stand in an entailment relation, the use of the logically weaker sentence systematically triggers the inference, called scalar implicature, that the logically stronger sentence isn't true. For instance, the sentence 'John ate a cookie or a muffin' triggers the inference that the logically stronger sentence 'John ate a cookie and a muffin' isn't true (Grice, 1975, a.o.). We observe that scalar implicatures triggered by quantified sentences with embedded disjunction (e.g. 'All of my students got As or Bs') will vary depending on the ratio between the cardinality of the restrictor (the number of relevant students) and the number of disjuncts. We argue that this effect can be explained if the relevant relation between sentences for implicature derivation is one of probabilistic informativeness rather than of logical entailment. Finally, in relation to (B), we demonstrate that the computation of probabilistic informativeness is performed

grammar-internally at least partly, because not all aspects of contextual knowledge enter into this computation.

In conclusion, this thesis uses novel experimental data and theoretical arguments to show two things. First, our findings add to a growing body of evidence that logical correlates of NPI licensing and scalar implicatures play a causal role in these phenomena. Second, our results suggest that logical correlates to NPIs are calculated post-grammatically, and logical correlates to scalar implicatures are calculated grammar-internally.

Résumé

Il a été démontré que plusieurs phénomènes linguistiques corrèlent avec certaines propriétés logiques de la phrase dans laquelle ils se trouvent, telles que les implications logiques de la phrase, ou les relations logiques entre la phrase et certaines phrases alternatives. Dans cette thèse, nous explorons (A) la question d'un éventuel rôle causal de ces propriétés logiques vis-à-vis de tels phénomènes linguistiques, et (B) la question de la manière dont ces propriétés logiques sont calculées. Il y a dans l'ensemble deux options concernant (B) : (i) ces propriétés logiques pourraient être calculées par un système formel n'ayant pas accès aux connaissances générales, appelons ce système formel la grammaire (Fox and Hackl, 2006, a.o.), ou bien (ii) ces propriétés logiques pourraient être calculées de manière post-grammaticale.

Le premier phénomène sur lequel nous nous concentrons dans cette thèse est celui de la légitimation des éléments à polarité négative (NPIs). Les NPIs sont des expressions telles que, par exemple, 'lift a finger' et 'anything' en anglais, qui sont acceptables dans certains environnements linguistiques ayant la propriété logique d'être "monotones décroissants" (Fauconnier, 1975a; Ladusaw, 1979). Un environnement linguistique est monotone décroissant s'il valide les implications logiques des ensembles (ex., 'tartes') vers les sous-ensembles (ex., 'tartes au potiron'). Dans le cadre de la question (A), nous nous intéressons à un phénomène particulier qui représente un défi pour la proposition selon laquelle les propriétés logiques de l'environnement ont un rôle causal pour la légitimation des NPIs (le phénomène en question est l'intervention dans la légitimation des NPIs). Nous examinons les prédictions expérimentales de la proposition selon laquelle l'intervention dans la légitimation des NPIs pourrait néanmoins être réduite aux propriétés logiques de l'environnement (Chierchia, 2004). Certains de nos résultats sont compatibles avec cette proposition, tandis que d'autres posent de nouvelles questions pour les théories de l'intervention dans la légitimation des NPIs. Pour ce qui est de la question (B), nous étudions la question de savoir si les NPIs exercent une influence sur les jugements de validité des inférences des ensembles vers les sous-ensembles et vice versa. Nos résultats expérimentaux démontrent que tel est bien le cas. Ces résultats, avec les résultats expérimentaux de Chemla et al. 2011, et les arguments empiriques présentés par Crnič 2014, suggèrent que les calculs post-grammaticaux des propriétés logiques des environnements jouent un rôle dans la légitimation des NPIs.

Le deuxième phénomène sur lequel nous nous concentrons dans cette thèse est celui des implicatures scalaires. Dans certains cas, quand une phrase X implique logiquement une autre phrase Y, l'usage de la phrase Y suggère que la phrase X est fausse. Par exemple, en entendant la phrase 'John a mangé un biscuit ou un muffin', on infère souvent que la phrase 'John a mangé un biscuit et un muffin' est fausse (Grice, 1975, a.o.). Cette inférence est une implicature scalaire. Nous observons que les implicatures scalaires des phrases quantifiées avec une disjonction telles que 'Tous mes étudiants ont eu un A ou un B' dépendent de la cardinalité du

restricteur du quantificateur (le nombre d'étudiants dans ce cas-ci) et du nombre de disjoints. Nous proposons une explication à cette observation dans le cadre où la relation pertinente entre phrases vis-à-vis de la dérivation des implicatures scalaires n'est pas celle des implications logiques, mais celle de l'informativité probabiliste.

Nous proposons que ce fait peut être expliqué si la relation entres phrases qui est pertinente pour la dérivation des implicatures scalaires n'est pas celle des implications logiques, mais celle de l'informativité probabiliste. En rapport avec la question (B), nous démontrons que certains aspects des connaissances générales ne sont pas accessibles au calcul de l'informativité probabiliste. Cela suggère que l'informativité probabiliste qui est pertinente pour les implicatures scalaires est calculée au niveau grammatical.

Pour conclure, cette thèse utilise de nouvelles données expérimentales et de nouveaux arguments théoriques pour démontrer deux choses. Premièrement, nos conclusions s'inscrivent dans la lignée d'un grand nombre de résultats attestant d'un rôle causal de propriétés logiques au sein de ces deux phénomènes. Deuxièmement, nos résultats suggèrent que les propriétés logiques corrélées aux NPIs font intervenir des calculs post-grammaticaux, tandis que les propriétés logiques intervenant dans les implicatures scalaires sont calculées au niveau grammatical.

Acknowledgments

I have spent the past six years at ENS, and it has been a very exciting and productive period, over which I learned immensely. The biggest thanks for that go to Emmanuel Chemla and Benjamin Spector. Emmanuel and Benjamin have always encouraged me to pursue my interests and helped me acquire the tools I needed to do so. I am extremely grateful for your constant availability and willingness to discuss virtually any question I was puzzling about, and that you were always generously offering your insights, advice, and support. You have been wonderful supervisors, and I am very happy to have had the opportunity to work with you.

I also want to thank David Barner, Gennaro Chierchia, Angelika Kratzer and Lyn Tieu for accepting to be members of my thesis committee, and for taking the time to think about my work and give me feedback on it. I have truly enjoyed being able to discuss my work with you.

Gennaro's work has often been an inspiration for mine, and I am very grateful to have had an opportunity to interact with him while I was spending a semester at Harvard. Gennaro, thank you so much for your interest in my work, and for all the discussions we had, which were always insightful, as well as for your kindness and encouragements.

Lyn co-supervised with Emmanuel the project on intervention effects and went with me step-by-step through my very first publication. She also taught me a lot on child language acquisition in the meantime. Lyn, thank you for all this — it was really a pleasure to work with you.

I would also like to thank a few other people who worked with me: Caterina Donati, who co-supervised my MA thesis project with Benjamin, Carlo Geraci and Vincent Homer, who supervised smaller research projects I undertook during my masters, and Yasu Sudo, who has been collaborating with me on a project on donkey anaphora, making the phenomenon look (somewhat) less scary and the work on it (much) more fun.

At ENS I have been a member of two labs: Laboratoire de Sciences Cognitives et Psycholinguistique, and Institut Jean-Nicod, in which I was a member of LINGUAE group. I would like to thank all of the members of the two labs for creating a nice working environment, for all the feedback they provided on my work, and for everything I learned from them.

I have been an inhabitant of the office 228 at 29 rue d'Ulm for four years, and I was lucky to share it at one time or another with a bunch of amazing people: in order of arrival, Mora Maldonado, Manuel Križ, Lyn Tieu, Amir Anvari, Émile Enguehard, Chloé Tahar, and Diego Feinmann. Thanks to you, every day in that office was a lot of fun, and every day I learned something new from you. Thank you all so much for your friendship throughout these years. Special thanks for everything we've been through together go to Amir Anvari, Mora Maldonado and Andrés Soria Ruiz — you have been such a great PhD cohort and the most wonderful friends one can hope for.

I am also grateful to many other current and former students and post-docs at ENS, not only for all the feedback they have given me on my work and for everything they taught me about their work, but also for their companionship in, among others, long coffee pauses, badminton matches, running tours around Paris and evenings out: Valentina Aristodemo, Mireille Babineau, Moshe E. Bar Lev, Morgan Beaurenaut, Giulio Bondanelli, Brian Buccola, Bianca Cepollaro, Alexandre Cremers, Anouk Dieuleveut, Flora Donati, Rachel Dudley, Grégory Dumont, Ewan Dunbar, Adriana Guevara Rukoz, Charlotte Hauser, Naomi Havron, Daniel Hoek, Jeremy Kuhn, Alice Latimier, Armando Lavalle, Allan Mancoo, Alex Martin, Gerda Ana Melnik, Cathal O'Madagain, Ignacio Rebollo, Louis Rouillé, Mirko Santoro, Camila Scaff, Martina Sundqvist, Tristan Thommen, Hualin Xiao.

In addition to everyone mentioned above, I would like to thank the following people for taking the time to discuss my work with me at various occasions: Lucas Champollion, Keny Chatain, Kathryn Davidson, Jakub Dotlačil, Danny Fox, Martin Hackl, Salvador Mascarenhas, Marie-Christine Meyer, Maša Močnik, Tom Roeper, Jacopo Romoli, Philippe Schlenker, Jesse Snedeker, Hedde Zeijlstra, Sarah Zobel.

I am lucky to know many fellow linguistics students around the world, who have been great companions through conferences, summer schools and travels: Hagen Blix, Paloma Jeretič, Jelena Kuzmanović, Karoliina Lohiniva, Daniel Margulis, Sergio Ramos, Giuseppe Ricciardi, Stanislao Zompi, and many many other friends from EGG and ESSLLI.

Before coming to ENS, I did my undergraduate studies in Ljubljana. I am grateful to Sašo Živanović, who was my undergrad supervisor: if it hadn't been for him, there's a good chance I would have ended up doing something else. In addition to Sašo, I would like to thank Gregor Perko and Dalibor Tomić for having been great teachers and for their help when I decided to apply to ENS.

I have moved around a lot, and this means that there are many people I care about who I only get to see rarely. Many thanks to my friends from Kruševac, Maribor, Ljubljana, and Paris, with whom I always feel that I am at home, whenever and wherever we get to see each other.

Moving around a lot also meant that I have been far away from my family, and that has been the hardest part of it. I am extremely grateful to my parents, Mile and Dragana, and to my sister Nadica, for always being there for me, and for all their love and support.

Finally, I am grateful to Marc, who has always been curious to hear about what I am working on, and has been filling my days with his wonderful insights on topics ranging from cognition and maths to animal life, series, and where to get the best brunches in Paris. Thank you for being by my side throughout these past few years, and for making them better in so many ways.

Contents

Part I

Thesis overview

Chapter 1

General introduction

Numerous linguistic phenomena have been shown to correlate with some logical properties of a sentence with which they occur, such as what logical entailments are supported by the sentence, or what the logical relation is between it and some other sentence.

One such phenomenon is the acceptability of negative polarity items (NPIs): NPIs are linguistic expressions such as *any*, whose acceptability has been linked to a class of logical entailments supported by the sentence in which the items occur (Fauconnier, 1975b; Ladusaw, 1979, a.o.). More specifically, NPIs are acceptable (= licensed) in downward-monotonic environments: these are environments in which inferences from supersets to subsets are valid. For instance, the NPI *any* is licensed in the sentence 'Harry didn't eat any pie': note that this sentence entails 'Harry didn't eat pumpkin pie' (in other words, the inference from superset to subset is valid). On the other hand, the NPI *any* is not licensed in the sentence 'Harry ate (*any) pie': note that this sentence does not entail 'Harry ate pumpkin pie' (in other words, the inference from superset to subset is not valid.)

Another such phenomenon is scalar implicatures: in certain cases, when two sentences stand in the entailment relation, the use of the logically weaker sentence systematically triggers the inference that the logically stronger sentence is false (for instance, the sentence 'John ate a cookie or a muffin' triggers the inference that the logically stronger 'John are a cookie and a muffin' is false) (Grice, 1975; Sauerland, 2004; van Rooij and Schulz, 2004; Schulz and van Rooij, 2006; Spector, 2006, 2007; Chierchia et al., 2012; Franke, 2011; Bergen et al., 2016, a.o.).

The existence of the correlation between logical properties of sentences and certain linguistic phenomena raises the question of whether these logical properties play a *causal role at the cognitive level* in these linguistic phenomena. By logical properties playing a causal role at the cognitive level we mean that the connection between logical properties and the linguistic phenomena in question is part of the linguistic competence of the speaker: the speaker needs to compute the logical properties in order to assess the relevant linguistic phenomena.

Option 1 is that logical properties of sentences do not play a causal role at the cognitive level in the linguistic phenomena in question: according to this view, the connection of linguistic phenomena with logical properties would be 'external' to speaker's competence (and would have to be explained differently, in terms of language evolution, for instance). To give a concrete example, in the case of NPIs, it has been argued that they are acceptable when a syntactic relation has been established between them and certain operators such as negation, which happen to create a downward-monotonic environment for the NPI (but it is the syntactic relation rather than the downward-monotonicity of the environment which plays a causal role in the licensing of NPIs according to this view) (Guerzoni, 2006; Szabolcsi, 2004, a.o).

Option 2 is that the logical correlates play a causal role at the cognitive level in these linguistic phenomena. If *Option 2* is on the right track, that opens up a major language and cognition architecture question: *at what level do calculations of these logical properties occur*?

There are two broad possibilities here: these logical properties could be computed grammar-internally, or these computations could be post-grammatical. Let us explain what we mean by this.

Numerous authors have argued, based on different linguistic phenomena, that there is a level of representation of the 'logical' meaning of a sentence, which is computed by a formal system which does not have access to contextual knowledge, nor to the meaning of lexical (roughly, open-class) categories (see Chierchia, 1984; Fox, 2000; Gajewski, 2002; Fox and Hackl, 2006; Chierchia, 2013 for arguments in favor of this view, as well as Dowty, 1979; Barwise and Cooper, 1981; Ladusaw, 1986; Von Fintel, 1993; Rullmann, 1995; Chierchia, 2004, 2006; Menendez-Benito, 2005 a.o. for relevant linguistic phenomena). In other words, this formal system computes the meaning of a sentence based on its 'logical' (roughly, closed-class) vocabulary, such as conjunctions, negation, quantifiers, prepositions etc. Let us call this formal system which does not have access to contextual knowledge or the meaning of lexical categories 'grammar'.

The first possibility is that the logical properties which have a causal role in the linguistic phenomena are calculated within this formal system, ie. within grammar. Let us call this possibility *Option 2a: grammar-internal*. For instance, in the case of NPIs, there could exist a mechanism in grammar which evaluates monotonicity properties of linguistic environments, and the result of this calculation plays a causal role in NPI licensing.

The second possibility is that the logical properties which have a causal role in the linguistic phenomena are calculated post-grammatically. Let us call this possibility *Option 2b: post-grammatical*. For instance, it could be that logical entailments of a sentence are evaluated outside of grammar, but that the result of this evaluation is then available to the procedure which determines the licensing of NPIs.

Deciding whether logical correlates of NPI licensing and of scalar implicatures, mostly entailment patterns and monotonicity, play a causal role in these linguistic phenomena, and if so, at what level these logical properties are calculated, is important both for our understanding of these specific linguistic phenomena, as well as more generally for our understanding of what computations can be performed grammar-internally, and what kind of post-grammatical computations affect language.

What kind of arguments can be put forward in relation to these questions?

Let us start with NPI licensing. There are two empirical arguments we are aware of in favor of *Option 2b: post-grammatical*.

The first one comes from Chemla et al., 2011. In this study, they ask how well subjective inferential judgments of monotonicity predict NPI acceptability judgments. These inferential judgments are naturally understood as *post-grammatical* computations: they might be affected by the specific predicates in the sentence and by different cognitive and contextual constraints etc. Note further that people are known to have difficulties *reporting* monotonicity properties of sentences (see for instance Geurts and van Der Slik, 2005). It is thus non-trivial that one's inferential judgments of monotonicity should predict one's NPI acceptability judgments. However, this is precisely what Chemla et al., 2011 find. In other words, Chemla et al., 2011 find that the

NPI judgments of a particular participant best correlate with this same participant's inferential judgments of monotonicity, suggesting that the computations involved in NPI licensing occur at the same, post-grammatical level as the computations participants use to produce their inferential judgments.

The second argument comes from the fact that contextual knowledge about predicates sometimes interferes with how acceptable an NPI is in a particular environment (cf. Crnič, 2014). Consider the minimal pair in (1): (1a) is reported to be better than (1b).

(1) a. Exactly two congressmen read any books last year.
b. ?Exactly two congressmen killed any people last year.

Crnič, 2014 argues that the differences between (1a) and (1b) can be explained if general world knowledge according to which congressmen are likely to read books and not likely to kill people somehow finds its way into NPI licensing.

Inferential judgments and contextual knowledge playing a role in NPI licensing can be taken as an argument that, if indeed the monotonicity of the environment plays a causal role in NPI licensing, it is calculated post-grammatically. Many challenges remain however both for the weak view according to which the monotonicity properties of the environment play a causal role in NPI licensing (*Option 2*) with no commitment to how these properties are calculated, and for the stronger view according to which these monotonicity properties are calculated post-grammatically (*Option 2b: post-grammatical*). Studies reported in Part II of this thesis address some of these challenges.

Let us now see what is known about scalar implicatures. It is uncontroversial that contextual knowledge plays some role in the implicature derivation: for instance, in certain contexts people derive fewer implicatures (Horn, 1972, a.o.). This however in itself does not demonstrate that the logical relation between sentences which feeds the implicature derivation is calculated post-grammatically. In fact, the data point in (2) discussed by Magri, 2009a can be taken in favor of *Option 2a: grammar-internal* calculation of this logical relation. Magri, 2009a observes that sentences such as (2) are deviant. He proposes that (2) is deviant because it has as implicature the negation of (3). This implicature, together with (2), contradicts common knowledge that all Italians come from the same country.

(2) #Some Italians come from a warm country.

(3) All Italians come from a warm country.

Crucially, given contextual knowledge, (2) and (3) are equivalent: (2) entails (3), and (3) entails (2). However, if (2) is indeed deviant because it triggers the implicature that (3) is false, then (3) needs to be considered by the implicature derivation mechanism as asymmetrically entailing (2). In other words, the mechanism which evaluates the logical relation between (2) and (3) cannot have access to contextual knowledge, suggesting that this logical relation is calculated grammar-internally.

In Part III of this thesis, we present novel empirical observations which we use as a further argument in favor of this view. These are important, as there have been alternative (non-implicature) accounts for the deviance of (2) (Spector, 2014b; Anvari, 2018). It is not clear as of yet whether these alternative accounts can be extended to capture the observations presented in Part III. Furthermore, in Part III we make the case that the relation between the two sentences which feeds the mechanism of implicature derivation is that of conditional likelihood, and not that of logical entailment. In other words, if the conclusions reached in Part III are on the right track, (i) the informativeness relation which feeds the mechanism for implicature derivation is computed in grammar; and (ii) this informativeness relation is that of conditional likelihood, and not that of logical entailment.

To summarize, this article-based thesis consists of work on the two linguistic phenomena introduced above: NPI licensing and scalar implicatures. These two phenomena are at the interface between linguistics and reasoning. They are used to study how logical properties can play a role in linguistic phenomena, and how these logical properties are computed so that they can make their way into linguistic judgments. This thesis thus contributes to the understanding of computations necessary for linguistic purposes, and to the evaluation of whether these computations are better thought of as happening within grammar, that is encapsulated from contextual knowledge, or post-grammatically. The studies reported in this thesis can be seen in Table 1.1.

Table 1.1: Studies reported in the thesis

	Part II: Polarity items
1	**Polarity items and reasoning** Submitted as: Denić, M., Homer, V., Rothschild, D., Chemla, E. (2019). The influence of polarity items on inferential judgments.
2	**Intervention effects in NPI licensing** Published as: Denić, M., Chemla, E., Tieu, L. (2018). Intervention effects in NPI licensing: A quantitative assessment of the scalar implicature explanation. *Glossa: A Journal of General Linguistics*, 3(1), 49
	Part III: Implicatures
1	**Blind ignorance inferences of embedded disjunction** To be submitted as: Denić, M. (2019). Blind ignorance inferences of embedded disjunction.
2	**Quantifier spreading in child language** To appear as: Denić, M., Chemla, E. (2019). Quantifier spreading in child language as distributive inferences. *Linguistic Inquiry*

Chapter 2

Overview of Part II

Monotonicity is an abstract, logical property that a linguistic environment is said to have when this environment systematically supports inferences from subsets to supersets (or *vice versa*). For instance, an environment is *upward monotone* or *upward-entailing* if it supports a subset to superset inference; an example is the environment of the boldface expressions in (1). Similarly, a *downward monotone* or *downward-entailing* environment supports the superset to subset inference; an example is in (2).

(1) This animal is **a siamese cat**.
$\rightsquigarrow$ This animal is **a cat**.

(2) This animal isn't **a cat**.
$\rightsquigarrow$ This animal isn't **a siamese cat**.

Interestingly, there is a class of expressions, called *polarity items* (PIs) whose acceptability has been linked to this logical property of monotonicity. This was first proposed by Fauconnier (1975b) and Ladusaw (1979), in relation to the most studied category of such expressions, namely negative polarity items (NPIs) such as *any*, *ever*, and *at all*. The generalization proposed for the distribution of NPIs is that they are felicitous in a downward-entailing environment, as in (3), and not felicitous in an upward-entailing environment, as in (4).

(3) This animal isn't a cat **at all**.

(4) *This animal is a cat **at all**.

In addition, there are *positive polarity items* (PPIs) such as *some*, *something*, *someone* that are acceptable in upward-entailing environments as in (5), but that cannot be interpreted in certain downward-entailing environments, such as the direct scope of unembedded negation. For instance, (6) doesn't have a reading in which the existential quantifier *some coffee* is interpreted with narrow scope with respect to negation (i.e. a reading equivalent to 'I didn't drink any coffee'). It is the lack of narrow scope of *some* under negation which signals that it is a PPI. Note that the sentence is acceptable under a reading where *some* outscopes negation (i.e. a reading equivalent to 'There is some coffee that I didn't drink'), possibly by moving covertly past negation, and therefore ending up being interpreted in an upward-entailing environment.

(5) I drank some coffee.

(6) I didn't drink some coffee. ($\neq$ I didn't drink any coffee.)

It should be noted, however, that NPIs are not a homogeneous group (different NPIs are acceptable in somewhat different sets of environments), and the same goes for PPIs (see Chierchia, 2013; Szabolcsi, 2004; van der Wouden, 1997; Spector, 2014a; Zwarts, 1998). Notwithstanding these complications, the simplified generalizations above will suffice as a starting point for interesting questions that can be asked about NPIs and PPIs.

The existence of NPIs and PPIs in language raises many questions. First of all, why would expressions with such complex distributional properties exist in language? In other words, what is the added value of using a polarity item (PI) instead of a semantically related expression that is not a PI? Also, what exactly is the connection between monotonicity and PIs? Does monotonicity play a causal role in their distribution? Also, how are these expressions acquired and processed? Do they relate in any way to (imperfect) subjective judgments of monotonicity, and can thus serve as a window into the interaction between language and reasoning capacities, or does polarity sensitivity rely solely on abstract logical monotonicity?

In the first part of the thesis, we present two projects addressing some of these questions by investigating the role of polarity items in reported inferences (see Chapter 5) and by studying so-called intervention effects (see Chapter 6). We summarize each of these projects below, and the complete projects can be found in relevant chapters.

2.1 Inferential judgments and polarity items licensing study

In the study reported in Chapter 5, we turn to the following question: how is the processing of polarity items related to subjective inferential judgments of monotonicity? This question is two-fold: (i) are subjective inferential judgments of monotonicity (as opposed to some abstract logical monotonicity) what predicts the intuitions of polarity items licensing, and (ii) do polarity items influence inferential judgments.

We present now the previous experimental work on these questions, which has generated somewhat conflicting results. The first relevant study is that of Chemla et al. (2011), which investigated the question (i): whether subjective monotonicity predicts polarity items licensing. They did this by collecting from a group of people both upward/downward inferential judgments and NPI acceptability judgments: it was found that the inferences a particular person considers valid in a given linguistic environment predict how acceptable they would find an NPI in that same environment. The generalization they reached is that the more someone finds downward inferences valid and upward inferences invalid in a given environment, the higher the acceptability rating this person will provide for an NPI in that environment. This study thus provided empirical confirmation of the relationship between monotonicity properties and NPI acceptability and suggests that *subjective* judgments of monotonicity are a better indicator of NPI acceptability than abstract logical monotonicity properties of the sentence.

Importantly, if the subjective judgments of monotonicity play a role in NPI licensing, one may expect that the presence of a polarity item in an environment should (in principle) be able to affect inferential judgments in that environment. The reason is that, to the extent that it is possible for the listener to process an NPI successfully in an environment, the listener needs to construe this environment as having the right subjective monotonicity properties. In other words, the presence of the NPI should increase the likelihood that the listener construes the environment in which the NPI is as having the monotonicity properties required for NPI licensing.

This question has been investigated in previous research, and no such effect was found. Szabolcsi et al., 2008 report a set of experiments well designed to prompt a potential facilitation effect of the presence of an NPI on corresponding inferences. They report on both explicit and implicit measures of inference facilitation (mere accuracy in inferential tasks, as well as reading times of phrases that presupposed the conclusion of a downward-entailing type of inference). They report no facilitation effect of the NPI.

This lack of result is challenging for the view according to which subjective inferential judgments mediate polarity items licensing.

In the study reported in Chapter 5, we take another look at this question. We hypothesize that if such an effect was to be found, it would not necessarily be found in environments where the monotonicity judgments are straightforward, because the polarity item has not much room to change anything. We thus focus on two environments in which the monotonicity judgments are more difficult: the non-monotonic environments, and the environments with two downward-entailing operators (which we refer to as double negative environments). Contrary to previous findings, we do find an effect of polarity items on inferential judgments.

NPIs. We find that negative polarity items ***in a non-monotonic environment*** make one perceive the non-monotonic environment as more downward-entailing as compared to when the negative polarity items are absent in that environment. In other words, people report more often that (7a) suggests (7b) and that (7b) does not suggest (7a) when the negative polarity item *any* is in the premise as compared to when no polarity item is in the premise.

(7) a. Exactly 12 aliens saw (any) birds.
b. Exactly 12 aliens saw (any) doves.

This effect is small but stable, as it is replicated multiple times through four experiments.

We also find some evidence that NPIs inflience monotonicity judgments ***in double-negative environments***: the meta-analysis of the two experiments in which we tested the inferences in this type of environment suggests that NPIs might make people perceive these environments as more downward entailing when the NPI is in the premise as compared to the baseline when no polarity item is in the premise. But this effect is not as stable as the effect of NPIs in non-monotonic environments. We discuss a number of possible reasons why this might be the case, including power considerations.

(8) a. Every alien who didn't see (any) birds is hairy.
b. Every alien who didn't see (any) doves is hairy.

PPIs. We find no evidence that PPIs have an effect on inferential judgments ***in non-monotonic environments***. In other words, in our experiments, we find no evidence that adding the positive polarity item *some* in sentences such as (9a,b) influence people's inferential judgments of monotonicity with those sentences.

(9) a. Exactly 12 aliens saw (some) birds.
b. Exactly 12 aliens saw (some) doves.

Interestingly however, we find some evidence for the effect of positive polarity items on inferential judgments ***in double-negative environments***. In other words, people seem to perceive (10a,b) as more upward-entailing when the PPI is in the premise as compared to the baseline when no polarity item is in the premise.

(10) a. Every alien who didn't see (some) birds is hairy.
b. Every alien who didn't see (some) doves is hairy.

This last result needs to be interpreted cautiously though in light of (i) the lack of effect of PPIs in non-monotonic environments; and (ii) the possibility of the inverse scope of PPIs, which might facilitate upward inferences for reasons independent to the polarity nature of the item.

Nonetheless, focusing on the stable effect of NPIs in non-monotonic environments, these results demonstrate that the presence of negative polarity items influences people's subjective monotonicity judgments. They thus eliminate the challenge posed by the results of Szabolcsi et al., 2008 for the view that the kind of monotonicity which matters for NPI licensing is the subjective monotonicity.

There is a further interesting consequence of these results, relating to the influence of language on reasoning that we have observed. Strikingly, linguistic expressions of a closed class which in themselves do not induce the monotonicity of an environment can nonetheless influence inferential monotonicity judgments, and thus serve to filter out potential misunderstandings. This might be the (functional) reason why polarity items exist and are so numerous both within and across languages (but see Barker, 2018 and Denić, 2015 for a different take on this question based on scope considerations).

Finally, we outline a few possible directions for future work.

1. If monotonicity of the environment is what plays a causal role in polarity items licensing, children arguably need to be able to notice the connection between monotonicity of the environment and the polarity items distribution in order to be able to acquire the semantics of polarity items. Noticing this connection is arguably not trivial, which puts in doubt that 4-5 year olds already have adult-like semantics of polarity items, as has been argued by Tieu and Lidz, 2016. From the point of view of child language acquisition, then, it would be interesting to see whether the influence of polarity items on inferential judgments obtains in this population. If the effect we obtain with adults is replicated in children, this would demonstrate that such young children indeed have all the tools necessary to grasp the semantics of polarity items, thus strengthening the conclusion by Tieu and Lidz, 2016.

2. One possible interpretation of our results from the second study is that the influence of polarity items on inferential judgments is not due to semantic requirements of polarity items that the environment have certain monotonicity properties, but due to people observing and learning the correlation between monotonicity properties of an environment and a particular linguistic expression (much like linguists did). In order to evaluate this, one could manipulate the amount of environments in which a participant sees a particular PI, and see whether this influences the effect of a PI on inferences (the idea being that the more environments the licensed PI appears in, the more likely the participant is to learn the correlation between the monotonicity of an environment and the PI, and thus the more likely the PI is to have an effect on inferences). Furthermore, one could look at polarity items which are less frequent in language in general (such as for instance the minimizer *lift a finger*), and see whether their effect on inferential judgments is smaller than the effect of more common polarity items such as *any* (the idea again being that the

more common an item is, the more likely the correlation between its distribution and the properties of environments which it appears in is to be learned).

3. It would be interesting to understand better the effect of NPIs in double negative environments. In particular, if indeed NPIs in double negative environments make people perceive those environments as more downward-entailing globally, this opens up a possibility that NPIs are licensed in double negative environments not because the local environment of the NPI is downward-entailing, but because people mis-perceive the global environment, which is logically upward-entailing, as subjectively downward-entailing.

4. We have mentioned that NPIs are not a homogeneous class, and that PPIs aren't either. Focusing on PPIs, it is conceivable that monotonicity of the environment plays a causal role in the licensing of some of them, but not others. Interesting monotonicity-based proposals have been put forward to account for the distribution of a subset of PPIs called global PPIs (Spector, 2014a). In light of this it would be informative to examine whether global PPIs influence monotonicity judgments more that the PPI *some* we tested.

2.2 Intervention effects study

Simplifying somewhat, there are two main approaches to negative polarity items licensing: following Homer, 2012, we will refer to them as environment-based and operator-based approaches. According to the environment-based approaches, the monotonicity of an environment is the determining factor in whether a polarity item is licensed or not (Chierchia, 2004; Gajewski, 2005; Homer, 2012, a.o.). Note that environment-based approaches typically incorporate some syntactic component too in order to accommodate that NPIs can be licensed in environments which are locally downward-entailing but globally upward-entailing, such as (11). According to the operator-based approaches, the monotonicity of an environment itself is not playing a causal role in NPI licensing: rather, it is the syntactic dependency with an operator such as negation (this dependency is typically thought of in terms of syntactic agreement) which determines whether a polarity item is licensed or not (Guerzoni, 2006; Szabolcsi, 2004, a.o). Operator-based approaches typically capture straightforwardly that NPIs can be licensed in sentences such as (11), because according to these approaches the presence of an operator rather than the monotoniciity of the environment determines whether an NPI is licensed or not.

(11) It's not the case that John didn't eat any pie.

Experimental studies which report connections between subjective monotonicity judgments and NPI licensing such as the studies reported in 5 and in Chemla et al., 2011 favor the environment-based approach (see also some arguments in favor of environment-based approaches in Homer, 2012). There is however an interesting phenomenon in NPI licensing which has received both operator-based and environment-based account, without a consensus on which of the two accounts is superior. The phenomenon in question is intervention effects in NPI licensing. We propose to contribute further to the debate between environment-based and operator-based approaches to NPI licensing by exploring experimental predictions of an environment-based approach to intervention effects.

When certain elements, so-called *interveners*, occur in between the licensor and the NPI, they give rise to so-called intervention effects: the negative polarity items become unlicensed due to an intervener despite appearing in a logically downward-entailing environment (Linebarger, 1987; Krifka, 1995; Kroch, 1974; Szabolcsi, 2004; Chierchia, 2004; Beck, 2006; Guerzoni, 2006). For example, the universally quantified NP is an intervener, but the definite NP is not: consider the two pairs in (12) and (13), which differ only in the intervener status of the indirect object. (12a) and (13a) entail (12b) and (13b), respectively, yet adding the NPI in (12a,b) leads to degradation, while adding the NPI in (13a,b) does not.

(12) a. Monkey didn't give every rabbit (*any) juice.
b. Monkey didn't give every rabbit (*any) strawberry juice.

(13) a. Monkey didn't give the rabbits any juice.
b. Monkey didn't give the rabbits any strawberry juice.

Another intervener is the conjunction *and*: when conjunction is in the scope of negation, (14a) entails (14b), yet the NPI *any* is not licensed.

(14) a. Ana didn't bake both cookies and (*any) muffins.
b. Ana didn't bake both cookies and (*any) chocolate muffins.

Strikingly, interveners form a natural class: they are items that trigger so-called scalar implicatures in DE environments (Chierchia, 2004). Roughly, scalar implicatures are optional inferences arising when (i) a sentence can be argued to have a minimally different *alternative*, obtained for instance by replacing some lexical item with a similar one; for instance, *some* and *all*, or *or* and *and* would count as similar, or *scale-mates* (see Horn, 1972; Gazdar, 1979, as well as Katzir, 2007 and Fox and Katzir, 2011 for recent discussions about the derivation of alternatives), and (ii) the sentence is consistent with the alternative being false (see Grice, 1975; Sauerland, 2004; van Rooij and Schulz, 2004; Schulz and van Rooij, 2006; Spector, 2006, 2007; Chierchia et al., 2012; Franke, 2011; Bergen et al., 2016 for refinements and discussions from a variety of perspectives). In these conditions, the negation of the alternative may be added to the meaning of the sentence, as a scalar implicature. This is best understood through an example. (15a) can be argued to have (15b) as an alternative, obtained by replacing the universal quantifier *every* with the existential quantifier *any*. Because the environment is DE, the alternative (15b) is logically stronger than the sentence (15a), and therefore the negation of the alternative (15c) is compatible with the sentence, hence it may be added as a scalar implicature of the sentence.

(15) a. Sentence: Monkey didn't give every rabbit juice.
b. Alternative: Monkey didn't give any rabbit juice.
c. Scalar implicature: Monkey gave some of the rabbits juice.

Given that interveners are items that trigger scalar implicatures in downward-entailing environments, a natural hypothesis, put forward in Chierchia, 2004, is that there is a link between the two. To make things slightly more concrete, one could say that scalar implicatures add a non-downward-entailing component to the meaning of a sentence, and for precisely this reason they may disrupt the licensing of an NPI. The proposal according to which implicatures are the cause of intervention effects because they change the monotonicity properties of the environment in which the NPI occurs can thus be viewed as an environment-based approach to intervention effects. As it has been mentioned previously, operator-based approach to intervention effects exists as well (Guerzoni, 2006): simplifying somewhat, according to operator-based approach to intervention effects, they arise because the intervener disrupts the syntactic dependency between the polarity item and the operator which licenses it.

In the study reported in Chapter 6, we explore the predictions of the environment-based approach. We try to assess for the first time whether implicatures can or cannot be directly connected to intervention effects, by manipulating and testing implicature judgments (and thus inferential judgments) on the one hand and acceptability judgments for intervention configurations on the other hand.

A major objection to the environment-based approach to intervention effects is that implicatures are in principle optional: we know that implicatures can be suppressed, and if implicatures are the cause of intervention effects, supressing implicatures should eliminate intervention effects. Yet, intervention effects are reported in the literature to appear uncancellable. In Experiment 1 reported in Chapter 6, we demonstrate that, contrary to common assumptions, NPIs in intervention configurations have an intermediate status: they are not as good as clear cases of licensed NPIs, and not as bad as clear cases of unlicensed NPIs, with a large between-speakers variation. This result has the potential to reconcile intervention effects with its implicature account: it is possible that this overall better status of NPIs in intervention configurations is precisely due to the possibility to suppress implicatures, and that the variation observed is due to differences among speakers in terms of the extent to which they are able to interpret a sentence without an implicature.

Experiments 2, 3, and 4 reported in the paper search for a direct link between one's individual tendency to derive an implicature, and one's perception of intervention effects, with the expectation that the more someone is willing to suppress an implicature, the better he should judge sentences with intervention effects. In these three experiments, we measure the participants' tendency to derive implicatures, and their perception of intervention effects in two separate tasks. Across three experiments, we do not observe a correlation between the two measures. This leaves us in a new state of puzzlement: there is indeed variation in the strength of intervention effects, but this variation remains now unexplained, since these experiments do not link it to variation in the implicature domain.

In Experiment 5, we approach the problem from a different angle, hypothesizing that the link between implicatures and intervention effects was missed in Experiments 2-4 because the two measures come from very different tasks (variants of truth value judgment tasks were used to measure participants' tendencies to derive implicatures, and acceptability task was used to measure their perception of intervention effects). It is therefore possible that one's tendency to derive implicatures in a truth-value judgment task does not predict one's tendency to derive implicatures in the acceptability judgment task, and thus cannot possibly predict intervention effects. Our goal in Experiment 5 is to approach intervention effects and implicature derivation rate *simultaneously*. We do this by means of a novel experimental design which we call *the repair strategy* design. In short, we test how people interpret sentences with intervention effects, such as (16). Assuming that most participants find the sentence (16) ungrammatical, in order to interpret it, they have to repair it somehow. The question of interest is whether suppression of an implicature is a possible repair strategy for sentences such as (16). We therefore propose a new paradigm in which we hope to gain valuable insight from the way participants would interpret what are otherwise ungrammatical sentences.

(16) *Monkey didn't give every rabbit any juice.

We find that when people are forced to assign an interpretation to a sentence such as (16), the interpretation they assign to it is significantly more often *the one without the implicature*, as compared to parallel sentences with the negative polarity item.

The result of Experiment 5 needs to be interpreted with care in light of the lack of results from Experiments 2-4. Nonetheless, this result suggests that implicatures might indeed matter for intervention effects.

Before concluding, let us discuss how our results connect to one specific version of the proposal according to which intervention effects are caused by scalar implicatures. Many authors have proposed that the requirement of the NPIs to appear in downward-entailing environments is due to their activation of alternatives which must be used up by a covert alternative-sensitive operator akin to *only* or *even*: this procedure delivers a consistent interpretation only in environments with the right monotonicity properties (Chierchia, 2006; Krifka, 1995; Lahiri, 1998; Crnič, 2014, a.o.). Likewise, there is an approach to scalar implicatures according to which scalar implicatures are derived when an alternative-sensitive operator akin to *only* is attached to a sentence (Chierchia, 2004, 2006; Fox, 2007; Chierchia et al., 2012). It is thus conceivable that the same alternative-sensitive operator is behind both the derivation of scalar implicatures and the licensing of NPIs, and that whenever an NPI is licensed, scalar implicatures are necessarily derived (cf. Chierchia, 2013 for a proposal along these lines). Under this instance of the proposal according to which intervention effects in NPI licensing are due to scalar implicatures, no correlation is predicted between one's individual tendency to derive an implicature and one's perception of intervention effects. This view is thus compatible with the lack of correlation from Experiments 2-4. It is however unclear how this view can accommodate the variation in the perception of intervention effects observed in Experiment 1, and the connection between scalar implicatures and intervention effects observed in Experiment 5.

We leave it for future work to distinguish between the more general version of the proposal according to which intervention effects in NPI licensing are perceived when scalar implicatures are derived, and a specific version of the proposal which adds to this that whenever NPIs are licensed, scalar implicatures are necessarily derived.

Finally, we outline some further directions for future work on the question.

1. Another way to approach the relation between implicatures and intervention effects might be to test preschool children: they have been argued to have adult-like semantics of NPIs (Tieu and Lidz, 2016), yet they seem to derive fewer scalar implicatures than adults (Chierchia et al., 2001; Foppolo et al., 2012; Guasti et al., 2005; Papafragou and Musolino, 2003; Noveck, 2001, a.o.). The prediction of the implicature approach is that children should be less sensitive to intervention effects in NPI licensing.

2. In our experiments, we have observed that there is variation in the strength of intervention effects between participants. Implicature is a natural explanation for it. However, if it turns out that implicatures are not what is behind intervention effects and thus cannot be the cause of variation and overall better acceptability judgments of NPIs in intervention effects environments, future work will have to trace the source of this variation and

improvement, as well as determine what else, in addition to monotonicity, plays a role in NPI licensing.

2.3 Conclusion of Part II

The first question we have explored in Part II is how polarity items licensing relates to subjective inferential judgments of monotonicity. We find that the presence of polarity items in an environment modifies subjective monotonicity judgments of that environment. Together with the results from Chemla et al., 2011, these results suggest that monotonicity indeed plays a causal role in polarity items licensing (at least in the case of NPIs), and furthermore, that the monotonicity which underlies polarity items licensing may be subjective rather than abstract logical monotonicity.

The second question is whether implicatures are behind intervention effects in NPI licensing: if it can be shown that they are, this would be a further important argument in favor of environment-based approaches to NPI licensing. We have found some suggestive evidence in this direction, but more research is needed on the question (see above).

Let us sum up the conclusions from Part II. To the extent that empirical results so far favors environment-based approaches to NPI licensing and that subjective monotonicity judgments play a role in NPI licensing, they are most naturally accomodated under the view that monotonicity plays a causal role in NPI licensing, and that this logical property is calculated post-grammatically (ie. they are most naturally accommodated under *Option 2b: post-grammatical* described in Chapter 1).

For completeness, we discuss here how one of the explanatory environment-based theories of polarity items licensing, that of Chierchia, 2006, fits with post-grammatical monotonicity judgments playing a role in this phenomenon. Note that we will present a simplified version of the account put forward in Chierchia, 2006 (see also Crnič, 2014 for a related account). Quantifiers are standardly assumed to quantify over a contextually supplied domain (Westerståhl, 1989). Kadmon and Landman, 1993a have proposed that polarity items like *any* are existential quantifiers over wider domains than plain existential quantifiers, such as the indefinite *a* (see also Krifka, 1995; Lahiri, 1998 in this connection). Chierchia, 2006 builds on this idea to explain monotonicity requirements of negative polarity items. More specifically, he proposes that the sentence (17a) activates alternatives such as (17b): any_D, any'_D stand for quantification over some domains D, D', with $D' \subset D$.

(17) a. John didn't eat any_D cookies.
b. John didn't eat $any_{D'}$ cookies.

The meaning expressed by these propositions is respectively in (18a,b):

(18) a. $\neg\exists x \in D\ [cookies(x) \wedge ate(John, x)]$
b. $\neg\exists x \in D'\ [cookies(x) \wedge ate(John, x)]$

By assumption, the alternatives activated by (17a) are necessarily used up by an operator akin to *even*: informally, the semantic import of this covert *even* is to assert the proposition it attaches and that it is less likely than its alternatives.

This means that (17a) is parsed as in (19a), and gets interpreted as (19b).

(19) a. *even* [John didn't eat any_D cookies.]
b. John didn't eat any_D cookies & 'John didn't eat any_D cookies' is less likely than 'John didn't eat $any_{D'}$ cookies'

Given that (17a) is logically stronger than (17b), the semantic import of the covert *even* is consistent.

The situation changes when *any* occurs in upward-entailing contexts: as $D' \subset D$, (20a) is logically weaker than (20b), and thus cannot be less likely. The semantic import of *even*, when it attaches to (20a), is thus inconsistent. This is why negative polarity items 'feel' ungrammatical in upward entailing environments, according to Chierchia, 2006.

(20) a. John ate any_D cookies.
b. John ate $any_{D'}$ cookies.

It is thus conceivable that the set of alternatives that the covert *even* operates on are compared post-grammatically in terms of their relative likelihood. It would then follow that the more someone considers an environment to be downward-entailing, the more likely they are to consider the semantic import of *even* consistent when attached to a sentence with an NPI.

Chapter 3

Overview of Part III

Logical meanings of sentences are often enriched by inferences about the speaker and their intentions. Scalar implicatures are one of such inferences: if a speaker utters (1), we might draw the implicature in (2).

(1) John ate some of the cookies.

(2) John didn't eat all of the cookies.

All theories of scalar implicatures rely on the existence of a set of alternative utterances that the speaker could have said instead of the original assertion. The hearer of an utterance makes use of these alternatives in their derivation of scalar implicatures.

One central question for theories of scalar implicatures is thus what the alternative set is for a given utterance. In other words, which criteria a sentence has to satisfy in order to qualify as an alternative to some utterance?

Another central question is what the nature of the mechanism which derives the implicatures is. Two main views on the question are opposed in the literature. According to the first view, these inferences are a result of a pragmatic process, in which the hearer engages in reasoning about the alternative utterances the speaker might have said and draws conclusions about the speaker's beliefs about these alternatives (Grice, 1975; Sauerland, 2004; van Rooij and Schulz, 2004; Franke, 2011; Bergen et al., 2016, a.o.). According to the second view, these inferences are semantic in nature: they are in fact part of the logical meaning of the sentence (Chierchia, 2004; Chierchia et al., 2012; Fox, 2007, a.o.).

The same questions can be asked about child language. What alternatives a sentence activates in child language? What mechanism derives implicatures? It is conceivable that there are differences between adult and child language along both dimensions.

In the second part of the thesis, we present two projects addressing some of these questions by studying ignorance inferences (see Chapter 7) and investigating the phenomenon of quantifier spreading in child language (see Chapter 8). We summarize each of these projects below, and the complete projects can be found in relevant chapters.

3.1 Blind ignorance inferences of embedded disjunction

3.1.1 Criteria on alternatives

Let us first introduce briefly two main approaches to implicature derivation: (i) the pragmatic (Gricean) approach, and (ii) the grammatical (non-pragmatic) approach.

Consider again the sentence (1), repeated here, which triggers the implicature in (2). Let us grant that the only alternative (1) activates is (3) (we will for the time being not motivate why (3) is an alternative to (1), nor why it would be the only one).

(1) John ate some of the cookies.

(2) John didn't eat all of the cookies.

(3) John ate all of the cookies.

According to the pragmatic approach to implicatures, the interlocutors assume each other to adhere to certain 'rules' in their conversational behavior; one such 'rule' is that they should disclose all of the relevant information they possess (Gricean maxim of quantity). These 'rules' allow interlocutors to draw inferences about each other's information state. For instance, note that (3) is logically stronger (and thus contains more information) than (1). According to the maxim of quantity, then, to the extent that the speaker believes (3), they should say (3) rather than (1). The hearer of (1) might thus reason that the speaker would have said (3) had they believed (3) to be true. If in addition to this the hearer believes that the speaker is opinionated about the truth of (3), the hearer will conclude that the speaker believes (3) to be false. The consequence of this reasoning is the implicature in (2).

According to the grammatical approach, implicatures are not the result of reasoning about what the speaker knows: instead, they are part of the meaning of the speaker's utterance due to a silent alternative-sensitive operator, called the exhaustification operator *exh*.

The semantics of this silent operator *exh* is given in (4), it is very similar to that of the focus operator *only* (Chierchia, 2004, 2006; Fox, 2007; Chierchia et al., 2012). In short, the semantic import of *exh* is to negate as many alternatives as possible from the alternative set.

(4) a. $Exh(p) = p \wedge \bigwedge_{q \in IE(p)} \neg q$
b. $IE(p) = \bigcap\{A' \subseteq ALT(p)$: A' is a maximal set in $ALT(p)$ which can be negated consistently with $p\}$

According to the grammatical approach to implicatures, (1) is in fact parsed as (5) when it triggers the implicature in (2).

(5) *exh*(John ate some of the cookies)

Given that we have assumed that the only alternative of (1) is (3), this alternative can be negated consistently with (1), and thus the logical meaning of (5) is in (6).

(6) John ate some of the cookies and he didn't eat all of the cookies.

To summarize, according to the two approaches, implicatures have very different status. According to the pragmatic approach, they are pragmatic inferences, the calculation of which might incorporate various aspects of contextual knowledge and beliefs about the speaker. According to the grammatical approach, implicatures are in fact part of the logical meaning of the sentence: once we know the set of alternatives and that the sentence is parsed with an *exh*, implicatures are incorporated into the meaning of the sentence due to the semantics of *exh*. This is not to say that in the grammatical theories the inferences about the speaker and their intentions cannot play any role in implicature derivation: for instance, there is still room for them to play a disambiguation role (whether the sentence is parsed with an *exh* or without), or the role in which alternatives are considered for implicature derivation. Whether or not they play such roles is an empirical question.

Let us now move to the question of alternatives: which alternatives a sentence activates?

A strong evidence that the set of alternatives needs to be constrained in some way is the so-called *symmetry problem*[1] (Fox, 2007; Fox and Katzir, 2011). The symmetry problem arises when a sentence activates two alternatives such that (i) the meaning of the sentence is equivalent to the disjunction of the two alternatives; (ii) the two alternatives contradict each other. Let us illustrate this on the case of (1).

(1) John ate some of the cookies.

It is conceivable that both (7a) and (7b) could be alternatives of (1).

(7) a. John ate some but not all of the cookies.
b. John ate all of the cookies.

Both (7a) and (7b) are logically stronger than (1). If both were indeed the alternatives of (1), no scalar implicature could be derived by either the pragmatic Gricean approach, or the grammatical approach. The ability of these two approaches to account for the scalar implicature of (1) thus crucially relies on (7a) not being an alternative of (1).

This fact has lead many authors to adopt various structural constraints on alternatives, which ensure that (7a) is not an alternative to (1), while (7b) is. An early solution to this were Horn scales: certain lexical items are scalar items which are by assumption members of a scale of items, and the alternatives a sentence activates are those obtained by the replacement of a scalar item by its scale-mates (Horn, 1972). Examples of such scales are {some, all}, {or, and}, etc. More recently, Katzir, 2007; Fox and Katzir, 2011 have proposed a formulation of a more principled structural constraint on alternatives which likewise guarantees that (7b) but not (7a) is an alternative of (1).

The upshot of these solutions is that not every sentence qualifies as an alternative: there are structural constraints on alternatives which result in the utterance being associated with a set of *formal alternatives (FA)*.

Another generally accepted criterion for alternatives is relevance: an alternative to a given utterance has to express a relevant proposition in the context of an utterance.

To see why a relevance-based constraint on alternatives is needed, consider (8). While (8) typically triggers the scalar implicature in (8a), in some contexts it can also trigger the stronger implicature in (8b) or (8c) (Horn, 1972; Katzir, 2014).

(8) John did some of the homework.
a. John didn't do all of the homework.
b. John didn't do most of the homework.
c. John didn't do much of the homework.

The fact that (8) doesn't always trigger the implicatures in (8b) and (8c) is usually taken as

[1] According to Fox (2007), the symmetry problem was first stated in its general form in the class-notes of Kai von Fintel and Irene Heim.

an argument that the formal alternatives of (8) 'John did most of the homework' and 'John did much of the homework' in some contexts do not qualify as alternatives of (8) due to relevance reasons.

To summarize, we have seen two constraints on alternatives: structural and relevance-based constraints. In other words, alternatives a sentence S activates $ALT(S)$ in a given context are all those formal alternatives which are relevant in that context (cf. (9)).

(9) **Alternatives of a sentence *S*:**
$ALT(S) = \{[\![X]\!] : X \in FA(S)\} \cap \{p : p \text{ is a contextually relevant proposition}\}$

In the study reported in Chapter 7, we argue for an additional criterion on alternatives: an alternative to a given utterance has to be sufficiently informative. An alternative of a given utterance is argued to be sufficiently informative if conditional probability of the literal meaning of the alternative given the literal meaning of the utterance is sufficiently low.

We argue for this additional criterion in light of novel observations about inferences of embedded disjunction. In particular, the domain size plays a role in which inferences are triggered by disjunction in quantified sentences.

To see this, compare (10) and (11). These two sentences differ only in the number of people in the domain of the universal quantifier (20 in (10) and 2 in (11)). Strikingly, however, they trigger very different inferences.

(10) is preferably interpreted with the so-called *distributive inferences* in (10a) (Chemla, 2009; Chemla and Spector, 2011; Crnič et al., 2015; Chierchia et al., 2012; Fox, 2007; Klinedinst, 2007; Spector, 2006, a.o), rather than with the so-called ignorance inferences in (10b). The situation seems to be quite the opposite when the domain size is small: (11) is preferably interpreted with ignorance inferences in (11b), rather than with the distributive inferences in (11a).

(10) **All 20 of Mary's friends are French or Spanish.**
   a. $\rightsquigarrow$ At least one of them is French.
      $\rightsquigarrow$ At least one of them is Spanish.
   b. $\not\rightsquigarrow$ The speaker is ignorant about whether or not all 20 of them are French.
      $\not\rightsquigarrow$ The speaker is ignorant about whether or not all 20 of them are Spanish.

(11) **Both of Mary's friends are French or Spanish.**
   a. $\not\rightsquigarrow$ At least one is French.
      $\not\rightsquigarrow$ At least one is Spanish.
   b. $\rightsquigarrow$ The speaker is ignorant about whether or not both are Spanish.
      $\rightsquigarrow$ The speaker is ignorant about whether or not both are French.

In Chapter 7, we demonstrate how this can be accounted for if one assumes that alternatives have to satisfy informativeness criterion described above. The key is in the alternaitves such as those in (12): if the sentences such as (10) and (11) activate the alternatives in (12), ignorance inferences are derived; if they do not, distributive inferences are derived (this is true under various approaches to implicatures, all the details are in Chapter 7). Note that under certain assumptions, the sentences in (12) are more informative when the domain only contains two individuals as opposed to when it contains 20. The proposal is thus that the sentences in (12) pass the informativeness threshold for alternatives of (11), but do not pass it for alternatives of

(10).

(12) Some of Mary's closest friends is Spanish, Some of Mary's closest friends is French

The proposal put forward in Chapter 7 is thus to add another criterion (in addition to structural and relevance criteria) for alternatives, as in (13).

(13) **Alternatives of a sentence *S*: proposal**
$ALT(S) = \{[\![X]\!] : X \in FA(S)\} \cap \{Y : Y$ is a contextually relevant proposition$\} \cap \{Z : Z$ is an informative proposition$\}$

3.1.2 The nature of the mechanism

As we have outlined above, there are two main families of approaches to implicatures: the approaches according to which implicatures are a pragmatic phenomenon, and the approaches according to which they are part of the logical meaning of the sentence.

There are however different types of implicatures (for instance, scalar and ignorance), and it is conceivable that they are derived by different mechanisms.

Let us first discuss the case of scalar implicatures. Even though the approach according to which scalar implicatures are a pragmatic phenomenon might be conceptually more appealing, there are a number of empirical arguments against it. We will review here only one such argument, which we will refer to as *the blindness argument* (Magri, 2009a). Consider (14).

(14) #Some Italians come from a warm country.

The sentence in (14) is deviant: Magri, 2009a has argued that (14) is deviant because it triggers the implicature that not all Italians come from from a warm country. This implicature, together with the assertion (14), contradicts common knowledge according to which all Italians come from the same country, hence the deviance of (14).

The reason why this is an argument against pragmatic approaches to scalar implicatures is that, in principle, if scalar implicatures are the result of reasoning about what the speaker knows, it is surprising that the inferences according to which the speaker's epistemic state contradict common knowledge are derived.

In the second part of Chapter 7, we discuss novel observations, for which we argue to extend the blindness argument to ignorance implicatures. In particular, we argue that the deviance of sentences such as (15) is due to blind ignorance implicatures which contradict common knowledge.

(15) #Each of these girls is Mary, Susan, or Jane.

If the blindness argument is indeed a good argument in favor of non-pragmatic approaches to implicatures, the deviance of (15) presents an empirical argument for the non-pragmatic approaches to ignorance implicatures.

Importantly, if the deviance of (15) is caused by ignorance implicatures, *(15) evidences that the calculation of informativeness of alternatives has to be blind to certain aspects of contextual knowledge (in other words, it has to be grammar-internal).* The rationale for this is the following: in order for the sentence (15) to trigger ignorance implicatures, alternatives such as 'Some of these girls are Mary' need to be considered for implicature derivation. However, given contextual knowledge, the probability that some of the girls is Mary given the literal meaning of (15) is 1! Hence, if the alternatives such as 'Some of these girls are Mary' pass the informativeness treshold, it has to be the case that contextual knowledge is not integrated in this calculation.

We outline here a few directions for future work.

1. The data discussed in Chapter 7 motivated proposing the threshold of informativeness criterion for alternatives of a given sentence. Whether this informativeness criterion can be detected in empirical phenomena other than the ones that motivated it remains to be seen.

2. We have put forward an empirical argument that ignorance implicatures are derived blindly to common knowledge. There is however one piece of contextual knowledge which clearly plays a role in implicature derivation: domain size. What pieces of contextual knowledge are accessible in implicature derivation and why is an open question.

3. We have seen some evidence for the grammar-internal informativeness computations in adult language: how do these grammar-internal computations develop? In other words, do we see evidence that these computations are not sensitive to contextual knowledge even in young children, or does this property emerge as the language experience accumulates?

4. A related interesting question for future research, which, to the best of our knowledge hasn't received much attention in the literature, is whether the arguments that have been put forward in adult language in favor of non-pragmatic approaches to implicatures (such as blindness argument, but there are others) can be replicated in child language. If they cannot, this might suggest that implicatures in child language are derived via a very different mechanism than in adult language.

3.2 Alternatives in child language: the case of quantifier spreading

Children are known to have different implicature behavior from adults: in the majority of cases, children seem to derive fewer scalar implicatures (Chierchia et al., 2001; Foppolo et al., 2012; Guasti et al., 2005; Papafragou and Musolino, 2003; Noveck, 2001, a.o.).

There is a lot of work aiming to understand why child and adult language differ in this respect. An emerging consensus seems to be that children have fewer alternatives than adults. More specifically, many researchers have argued that children have difficulties accessing alternatives which require lexical replacements (cf. Barner and Bachrach 2010, Barner et al. 2011, Tieu et al. 2016, Singh et al. 2016, Tieu et al. 2017, and Pagliarini et al. 2018). For instance, when children hear (16a), they would have difficulties accessing the alternative in (16b), and hence would not be able to infer that (16b) is false.

(16) a. John ate some of the cookies.
b. John ate all of the cookies.

In the study B.2 reported in Chapter 8, we discuss a child language phenomenon called quantifier spreading. When an indefinite is in the scope of a universal quantifier, such as in (17a), children often interpret is as in (17b), as revealed by a large body of experimental work (Inhelder and Piaget, 1964; Roeper and de Villiers, 1991; Philip, 1995; Drozd and van Loosbroek, 1998; Geurts, 2003; Crain et al., 1996; Freeman et al., 1982; Brooks and Sekerina, 2006, a.o.). We propose that this phenomenon might reveal another difference in the implicature behavior of children and adults.

(17) a. Every girl saw a boy.
  b. Every girl saw a boy, and every boy (in some relevant domain) was seen by a girl.

We first demonstrate that, given standard assumptions, it is in fact surprising that adults do not interpret (17a) as (17b). Indefinites and disjunctions are often assumed to activate essentially the same alternatives in adult language (e.g. Chierchia 2013): everything else being equal, then, they should trigger the same inferences. This is however not the case. Consider (18a) uttered in the context in which the only boys are John, Bob, and Bill. (18a) triggers the distributive inference in (18b). Crucially, (17a), in which the disjunction is replaced by the indefinite, does not trigger the distributive inference in (17b): note that if it did, the adults would be exhibiting quantifier-spreading behavior with sentences such as (17a).

(18) a. Every girl saw John, Bill, or Bob.
  b. Some girl saw Bob, some girl saw Bill, and some girl saw John.

This opens up two somewhat independent research questions. The first one is why adults don't have distributive inferences with indefinites (=quantifier spreading)? There are two possibilities here: (i) the usual assumption that indefinites and disjunctions activate the same alternatives is wrong, (ii) indefinites and disjunctions activate the same alternatives, but there is an additional piece of grammar to be identified, which eliminates distributive inferences with indefinites. In line with (ii), we discuss two possible pieces of grammar which might be responsible for the difference between indefinites and disjunctions. Neither of them however captures the full range of available facts; we leave that question open for future research.

The second research question is whether quantifier spreading in child language is in fact a result of distributive inferences with indefinites, which are for reasons to be understood absent from adult language.

In response to this second question, we present in Chapter 8 an empirical argument which connects well with the hypothesis that quantifier spreading is a result of distributive inferences. Namely, it is known that certain contextual manipulations, such as increased contextual salience of the subject noun, decreases the rate of quantifier spreading: we demonstrate in an experimental setting that similar contextual manipulation decreases the rate of distributive inferences with disjunction in adults.

We outline here a few directions for future research.

1. If indeed quantifier spreading is due to distributive inferences of indefinites, this phenomenon is another instance of different implicature behavior of children and adults. What would be left to understand is then what exactly is different in child language: do indefinites activate the same alternatives as disjunction in child language, but not in adult language? Or do indefinites activate the same alternatives as disjunction in both child and adult language, but adult language contains an extra piece of grammar that children need to acquire, which eliminates distributive inferences with indefinites?

2. The proposal that quantifier spreading are distributive inferences with indefinites calls for

more research comparing the interpretation of indefinites in child language and disjunctions in adult language. The general prediction is that children should have quantifier spreading behavior whenever adults have distributive inferences with disjunction. One possible test case for the proposal is the comparison between what children and adults do with indefinites and disjunctions respectively in the scope of *all but one*. This test case is particularly interesting in light of alternative approaches to quantifier spreading, which make very different predictions from the distributive inferences approach.

3. Furthermore, the case of quantifier spreading and the possibility that it is an implicature phenomenon raises the question of whether there are other child language interpretation 'errors' that are in fact miscategorized linguistic inferences.

3.3 Conclusions of Part III

In the second part of the thesis, we explore three related questions: (1) what criteria alternatives of a given utterance have to satisfy; (2) what is the nature of the implicature derivation, focussing on the case of ignorance implicatures, and (3) what differences there are between implicature derivation, and in particular *alternative* derivation, in child and adult language.

In relation to question (1), we argue, based on novel empirical observations, that a threshold of informativeness is one of the criteria alternatives have to satisfy in order to feed the mechanism of implicature derivation. Interestingly, this criterion is most naturally phrased in terms of probabilities, rather than in terms of entailment relations, opening connections between the usual logical approach and probabilistic approaches to semantics and pragmatics. In relation to question (2), we offer novel empirical observations suggesting that *ignorance* implicatures may not be the result of pragmatic reasoning. In relation to question (3), we propose a novel account of quantifier spreading in child language, according to which this child language phenomenon is due to distributive inferences of indefinites in child language. If this proposal is on the right track, it reveals another difference between the implicatures in child and adult language. Two options are to be explored to explain the difference: (i) indefinites and disjunctions activate the same alternatives in child language but not in adult language; (ii) there is an extra piece in adult language which eliminates distributive inferences of indefinites that children who exhibit quantifier spreading have to acquire.

Focusing on question (1), we have seen evidence that the informativeness calculation does not seem to have access to contextual knowledge. This suggests that the notion of informativeness which is relevant for implicature derivation is calculated grammar-internally (cf. *Option 2a: grammar-internal* in Chapter 1).

Chapter 4

General conclusion

Let us now take a step back, and see what can be concluded about the role of logical correlates in the linguistic phenomena studies in this thesis. As a reminder, we have outlined in Chapter 1 a number of possible relations between linguistic phenomena and logical properties they correlate with. *Option 1* was that logical correlates do not play a causal role in these linguistic phenomena. *Option 2* was that that they do. If *Option 2* is on the right track, two possibilities are to be explored in terms of the level at which these logical properties are computed: *grammar-internally* or *post-grammatically*.

Let us start by discussing which of these options can be argued for in relation to negative polarity items licensing and their logical correlates.

There has been some empirical evidence in favor of the post-grammatical calculation of logical correlates of NPI licensing (Chemla et al., 2011; Crnič, 2014). This view however faces two types of challenges. First, previous experimental work did not find an effect of polarity items on inferential judgments of monotonicity (Szabolcsi et al., 2008). Such an effect would be expected though if indeed these judgments matter for polarity items licensing: in order for a person to succeed in interpreting a sentence with a polarity item, they need to construe the environment as having the right kind of monotonicity properties. The presence of a polarity item should thus increase the likelihood that the listener perceives an environment as having certain monotonicity properties. Second, there are certain phenomena in polarity items licensing, such as intervention effects, for which there is no consensus as of yet if monotonicity considerations can explain them. Their existence leaves the question of how big of a role the monotonicity plays in polarity items licensing open.

The studies reported in Part II of the thesis speak to these challenges.

The first study reported in Part II demonstrates that the presence of polarity items in sentences alters inferential judgments of monotonicity, and discusses why this effect was missed in the previous study that investigated this question.

The second study reported in Part II focused on intervention effects in NPI licensing. Both environment-based approaches and operator-based approaches to this phenomenon exist. We investigated predictions of an account according to which the monotonicity properties of the environment are behind this phenomenon (Chierchia, 2004). We obtained somewhat mixed results; the results of Experiment 5 of the intervention effects study however are suggestive of the role of monotonicity in this phenomenon, and they offer a promising avenue for further exploration of intervention effects in light of environment-based theories.

These two studies thus do away with some of the challenges to the idea that negative polarity items licensing is sustained by subjective inferential judgments of monotonicity.

To summarize, then, when it comes to negative polarity items licensing, *there is growing evidence that monotonicity indeed plays a causal role, and that the polarity items acceptability is sustained by post-grammatical monotonicity calculations (cf. Option 2b: post-grammatical in Chapter 1).*

Let us now move to scalar implicatures, and to how their logical correlates are calculated.

Traditionally, one of the factors that is assumed to play a role in the scalar implicature derivation is the logical relation between a sentence and its alternatives. Most theories of scalar implicatures give a causal role to this relation (*Option 2*). We have argued that the notion of informativeness which feeds the implicature derivation mechanism is that of conditional likelihood of an alternative given the utterance, rather than that of logical entailment of the utterance by an alternative, so it would be more precise to speak of 'conditional likelihood' correlates rather than logical correlates of scalar implicature derivation. Strikingly, however, there are empirical arguments that this calculation is blind to contextual knowledge: this suggests that this informativeness measure is not calculated by post-grammmatical mechanisms, which should presumably have access to contextual knowledge.

To the extent that this is on the right track, *there is a grammar-internal way to calculate comparative informativeness as measured by conditional likelihood between propositions, and it is utilized in the comparison of alternatives in scalar implicature derivation process (cf. Option 2a: grammar-internal in Section 1).*

The picture that is emerging is that NPI licensing and scalar implicature computation differ substantially in terms of how the logical correlates of these two phenomena are calculated (somewhat counter-intuitively, in a post-grammatical manner for the case of NPIs and grammar-internally for scalar implicatures). Taken at face value, this suggests that it's not the case that the exact same mechanism computes both scalar implicature computation and negative polarity items licensing (for a proposal along these lines, see (Chierchia, 2013)). This picture is however still compatible with very similar computations being behind the two phenomena, but with these performed at different levels.

To conclude, some of the findings of the studies reported in this thesis bring novel data that hopefully advances our understanding of how logical properties relevant for negative polarity items licensing and implicature derivation are calculated. Furthermore, the studies also contribute to our understanding of more specific questions in relation to polarity items licensing and implicatures, such as for instance what causes intervention effects in polarity items licensing, do polarity items influence inferential judgments, whether entailments or probabilistic informativeness matter for implicature derivation, and what differences there might be between child and adult language in terms of implicature derivation.

Part II

Polarity items

Chapter 5

Polarity and inferential judgments

This chapter is a reproduction of the following article:

Denić, Milica, Vincent Homer, Daniel Rothschild, and Emmanuel Chemla. The influence of polarity items on inferential judgments. Submitted.

Abstract

Polarity sensitive items are linguistic expressions such as *any*, *at all*, *some*, which are felicitous in some linguistic environments but not others. Crucially, whether a polarity item is felicitous in a given environment is argued to depend on the inferences (in the reasoning sense) that this environment allows. We show that the inferential judgments reported for a given environment are modified in the presence of polarity sensitive items. Hence, there is a two-way influence between linguistic and reasoning abilities: the linguistic acceptability of polarity items is dependent on reasoning facts and, conversely, reasoning judgments can be altered by the mere addition of seemingly innocuous polarity items.

5.1 Polarity and monotonicity

Monotonicity is an abstract, logical property that a linguistic environment is said to have when this environment systematically supports inferences from subsets to supersets (or *vice versa*). For instance, an environment is *upward monotone* or *upward-entailing* if it supports a subset to superset inference; an example is the environment of the boldface expressions in (1). Similarly, a *downward monotone* or *downward-entailing* environment supports the superset to subset inference; an example is in (2).

(1) This animal is **a siamese cat**.
$\rightsquigarrow$ This animal is **a cat**.

(2) This animal isn't **a cat**.
$\rightsquigarrow$ This animal isn't **a siamese cat**.

Interestingly, there is a class of expressions, called *polarity items* (PIs) whose acceptability has been linked to this logical property of monotonicity. This was first proposed by Fauconnier (1975b) and Ladusaw (1979), in relation to the most studied category of such expressions, namely negative polarity items (NPIs) such as *any*, *ever*, and *at all*. The generalization proposed for the distribution of NPIs is that they are felicitous in a downward-entailing environment, as in (3), and not felicitous in an upward-entailing environment, as in (4).

(3) This animal isn't a cat **at all**.

(4) *This animal is a cat **at all**.

In addition, there are *positive polarity items* (PPIs) such as *some*, *something*, *someone* that are acceptable in upward-entailing environments as in (5), but that cannot be interpreted in a number of downward-entailing environments. For instance, (6) doesn't have a reading in which the existential quantifier *some coffee* is interpreted with narrow scope with respect to negation (i.e. a reading equivalent to 'I didn't drink any coffee'). It is the lack of narrow scope of *some* under negation which signals that it is a PPI. Note that the sentence is acceptable under a reading where *some* outscopes negation (i.e. a reading equivalent to 'There is some coffee that I didn't drink'), possibly by moving covertly past negation, and therefore ending up being interpreted in an upward-entailing environment.

(5) I drank some coffee.

(6) I didn't drink some coffee. ($\neq$ I didn't drink any coffee.)

There is a large and complex literature seeking to refine and explain the exact conditions under which PIs are licit. Be that as it may, everyone agrees that there is some connection between the acceptability of NPIs and PPIs and the logical property of monotonicity: some authors argue that PIs are essentially connected with monotonicity, other may argue that monotonicity is not essential to the licensing constraints, and that it is, in the worst case, an accidental correlation. The former approach typically seeks reasons coming from the meaning of PIs that make them unacceptable outside of environments with certain monotonicity properties (Chierchia, 2013; Kadmon and Landman, 1993b; Krifka, 1991, 1995; Lahiri, 1998). The latter approach, instead, proposes that PI licensing is a form of syntactic agreement with some operators (e.g., negations), and that these operators happen to induce an environment of a particular monotonicity (see Guerzoni, 2006; Herburger and Mauck, 2007; Szabolcsi, 2004; Progovac, 2000, among others). Crucially though, all must recognize that this correlation between the monotonicity properties of the environment and the PI licensing exists. Even if it is an accidental correlation, it could be taken advantage of: PIs may serve as a signal of monotonicity, and their mere presence at the surface level could (even if probabilistically) facilitate corresponding monotonicity inferences.

In this paper we explore the relationship between polarity items and reasoning capacities. In particular, we explore the question of whether the presence of a polarity item influences what inferences subjects are willing to make. Szabolcsi et al. (2008) investigated this question before in relation to negative polarity items, concluding that there is no influence of NPIs on downward inferences. Given that there is at the very least a probabilistic link between PIs and monotonicity, it is surprising not to find such an effect. In the experiments reported below, we find such an effect in various cases, albeit not the cases tested by Szabolcsi et al., for reasons we will discuss. Thus, we demonstrate that the presence or absence of PIs in a sentence *does* influence which inferences subjects are willing to make, thereby demonstrating that high level reasoning tasks can be influenced by what otherwise looks like innocent linguistic decorations.

5.2 Previous results

Psycholinguistic studies have investigated the licensing of polarity items using a variety of tasks including acceptability judgments (e.g., Drenhaus et al., 2005; Muller and Phillips, 2018), ERP measures (e.g., Drenhaus et al., 2006; Drenhaus, Heiner et al., 2007; Saddy et al., 2004; Steinhauer et al., 2010; Shao and Neville, 1998; Yurchenko et al., 2013; Yanilmaz and Drury, 2018; Xiang et al., 2009), self-paced reading (e.g., Parker and Phillips, 2016; Xiang et al., 2013) and eye-tracking (e.g., Vasishth et al., 2008). Here we focus on two studies which jointly investigated the licensing of polarity items and inferential judgments of monotonicity: Chemla et al. (2011) and Szabolcsi et al. (2008).

Chemla et al. (2011) collected from a group of people both upward/downward inferential judgments and NPI acceptability judgments: it was found that the inferences a particular person considers valid in a given linguistic environment predict how acceptable they would find an NPI in that same environment. This study thus provided empirical confirmation of the relationship between monotonicity properties and NPI acceptability. In fact, these results also suggest that *subjective* individual judgments of inferential properties are a better indicator of PI acceptability than objective, logical upward and downward-entailingness. As in the experiments below, the study did not test all-or-nothing judgments of either NPI-acceptability or monotonicity inferences, but rather looked at graded judgments. The generalization reached about the determinants of NPI acceptability were more 'graded' than those in the syntax/semantics literature. In particular, they found that downward-entailingness and upward-entailingness *together* were a better predictor of NPI acceptability than either alone was, NPIs are thus good in environments to the extent that they are perceived as DE and/or as not-UE.

The other relevant study on the connection between PI acceptability in an environment and the monotonicity properties of that environment is the aforementioned Szabolcsi et al. (2008). They report a set of experiments well designed to prompt a potential facilitation effect of the presence of an NPI on corresponding inferences. They report on both explicit and implicit measures of inference facilitation (mere accuracy in inferential tasks, as well as reading times of phrases that presupposed the conclusion of a downward-entailing type of inference). They report no facilitation effect of the NPI.

In the experiments below we take another look at the question of whether PIs affect monotonicity judgments. Contrary to Szabolcsi et al. (2008), we show that PIs do in fact influence judgments of monotone inferences, just that these effects are (1) only present in cases in which the inferential patterns are less clear to subjects (that is, not in the most basic simple upward or downward-entailing environments); (2) they are not present for all polarity items in all tested configurations; and (3) they push participants more towards rejecting the incorrect inference rather than towards accepting the correct one.

The experimental material, data, the R script used for analysis, as well as the document with the output of all of the models reported in the paper can be found at https://semanticsarchive.net/Archive/WY4OTMzO.

5.3 Experiment 1: PIs affect the perception of monotonicity

Some environments give rise to clear (and correct) judgments of monotonicity: it is quite easy to see that 'John read a novel' entails that 'John read a book'. In such cases, adding a PI may not make the inferences any clearer, or lead people to change their mind in any way about it. In this experiment, we thus looked for an effect on inferences of PIs in contexts in which the inferential patterns are less clear. *Non-monotonic* environments do not support either subset to superset or superset to subset inferences (cf. ((7)); neither ((7a)) entails ((7b)), nor ((7b)) entails ((7a))). However, it has previously been shown that monotonicity judgments of these environments could be more graded, with participants reporting to a non-negligeable extent some monotonicity in one direction or another (see Chemla et al., 2011). Given this level of uncertainty as to whether these environments support upward or downward inferences, we hypothesized that the presence or absence of a PI might then have more room to influence the judgement. Importantly for our purposes, both PPIs and NPIs are known to be acceptable at least to a certain extent in these environments: both ((8a)) and ((8b)) can be interpreted as ((7a)) (there is however some individual variation in terms of NPI acceptability in NM environments, cf. Rothschild 2006; Crnič 2014; Chemla et al. 2011; Denić et al. 2018).

(7) a. Exactly 12 aliens saw birds.
b. Exactly 12 aliens saw doves.

(8) a. Exactly 12 aliens saw some birds.
b. Exactly 12 aliens saw any birds.

5.3.1 Method

5.3.1.1 Instructions and task

At the beginning of the task, the participants read the following instructions:

(9) *You will see pairs of sentences about aliens, who just spent last week on Earth. Imagine that you hear the first sentence, and indicate whether you would then naturally conclude that the second sentence is true.*

They were then given three examples of such pairs, call them premise-conclusion pairs. In one pair, the conclusion clearly followed from the premise ((10)), in a second one the conclusion clearly did not follow from the premise ((11)), and the third case was less clear ((12)).

(10) 'Each alien received a high score in all human IQ tests.' → Aliens are very intelligent.

(11) 'Few aliens visited Paris.' → All aliens visited the Eiffel Tower.

(12) 'Pink aliens have scary teeth.' → Pink aliens are the most terrifying.

The participants were instructed to record their responses on a continuous scale presented in the form of a bar by filling a portion of it red. They were explained that they could use the flexibility of the red bar to report intermediate judgments, and that they would get used to it naturally. The dependent measure was the percentage of the bar filled in red. This measure will be referred to as the 'rating' given to an inference.

5.3.1.2 Material

The material was made of pairs of sentences, which were intended to serve as the premise and the conclusion in some inferential judgment task. These pairs of sentences were constructed from the recombination of more atomic building blocks. Crucially, among these pairs there were both valid and invalid upward and downward inferences, with and without polarity items.

The building blocks to create these inferences were as follows. First, we created a list of 11 environments: 3 Upward Entailing (UE) environments (positive, Every, Many), 3 Downward Entailing (DE) environments (negative, No, Few) and 2 Non-Monotonic (NM) environments (Exactly 12, Only 12). Second, we created a list of 12 pairs of (superset, subset) VPs that could host a PI (see <PI> birds, see <PI> doves). We combined these two building blocks, environments and pairs of VPs, to obtain pairs of sentences for our inferential stimuli. Both orders of the pairs were used. Note that only one of the orders provides a valid inference in DE and UE environments, and neither of the orders provides valid inferences in NM environments.

Finally, for each of these pairs of sentences, we created items for which, in the premise, there was (i) no PI (for all environments), (ii) an NPI for DE and NM environments, (iii) a PPI for UE and NM environments. These possibilities correspond to all possibilities that may not be outrageously infelicitous (see discussions about the marginal acceptability of some PIs in NM environments in Rothschild 2006; Crnič 2014, and a quantitative evaluation in Chemla et al., 2011 and Denić et al., 2018).

Overall, we obtained 2 [superset/subset vs subset/superset] x 12 [VPs] x (3 [UE] x 2 [PPI vs no PI] + 3 [DE] x 2 [NPI vs no PI] + 2 [NM] x 3 [NPI vs PI vs no PI]) = 432 inference pairs. One example pair for each environment is provided in ((13))-((20)).

(13) Condition: UE-positive, superset → subset, (PPI)
   a. The purple alien saw (some) birds.
   b. The purple alien saw doves.

(14) Condition: UE-every, superset → subset, (PPI)
   a. Every alien saw (some) birds.
   b. Every alien saw doves.

(15) Condition: UE-many, superset → subset, (PPI)
   a. Many aliens saw (some) birds.
   b. Many aliens saw doves.

(16) Condition: DE-negative, superset → subset, (NPI)
   a. The purple alien didn't see (any) birds.
   b. The purple alien didn't see doves.

(17) Condition: DE-no, superset → subset, (NPI)
   a. No alien saw (any) birds.
   b. No alien saw doves.

(18) Condition: DE-few, superset → subset, (NPI)
   a. Few aliens saw (any) birds.
   b. Few aliens saw doves.

(19) Condition: NM-exactly 12, superset → subset, (PPI/NPI)
   a. Exactly 12 aliens saw (some/any) birds.
   b. Exactly 12 aliens saw doves.

(20) Condition: NM-only 12, superset → subset, (PPI/NPI)
   a. Only 12 aliens saw (some/any) birds.

b. Only 12 aliens saw doves.

These 432 items were distributed in three groups of 144 items each, so that: (a) all 12 VPs would appear in a group, (b) four different VPs were used in items which had a PPI in the premise, four different VPs were used in items which had an NPI in the premise, and four different VPs were used in items which had no PI in the premise, (c) across groups, all 12 VPs would appear with the three types of items (an NPI in the premise, a PPI in the premise, no PI in the premise). Hence, in each group there were 4 [items with different VPs] x 2 [set/subset vs subset/set] x (3 [UE] x 2 [PPI vs no PI] + 3 [DE] x 2 [NPI vs no PI] + 2 [NM] x 3 [NPI vs PI vs no PI]). Participants were administered one of these groups of items, presented each time in a random order.

Apart from the 144 target items, the participants in each group also had to provide responses to the three training items which were administered at the beginning of the task. They were identical to the examples discussed in the instruction, and their purpose was to let participants get used to the setting and to the task.

5.3.1.3 Participants and exclusion criteria

75 participants were recruited through Amazon Mechanical Turk (38 females). As the result of the following two exclusion criteria, the responses of 66 participants among them were kept for the analysis (32 females). First, the results of one participant were excluded for them reporting not being a native speaker of English. Second, the results from eight more participants were excluded for not judging downward inferences higher in downward than in upward monotonic environments, or for not judging upward inferences higher in upward than in downward monotonic environments. The rationale for this second exclusion criterion is that, as it is likely that these judgments should be straightforward and maximally polarized, these eight participants did not understand the task in the way we expected (or were responding at random). The same exclusion criteria were applied in all four experiments reported in this paper.

5.3.2 Results

Responses given in less than 1.4 s (1% of the data) or more than 10s (9% of the data) were removed from the analysis. These numbers were chosen by a visual inspection of the distribution of RTs, with the goal of removing clear outliers. We excluded more responses falling on the slow part of the spectrum, to exclude non-spontaneous responses. This criterion was thus chosen by hand, looking only at the RTs and not the condition and responses they corresponded too. It was then copied without change for the following experiments, which provide replications of these results.

The three training items were answered as expected, with ratings of 93% for the clearly valid inference (10), 9% for the clearly not valid one (11), and 68% for the intermediate one (12).

5.1 represents on the y-axis ratings of inferences with subset in premise and superset in conclusion. These inferences were valid for UE environments. These ratings thus measure the perceived UE-ness of the environment, and we refer to them as UE-ratings. On the x-axis, the ratings correspond to superset to subset inferences, which were valid in DE environments, and are accordingly referred to as DE-ratings. The graph then reports the mean such rating across participants and across the three types of environments (UE, DE, and NM). The graph shows that participants were behaving properly on these broad distinctions: disregarding the effect of PIs for the time being, UE environments ended up in the top left corner of the graph with high

UE-ratings (89.7%, $SD = 13.03$) and low DE-ratings (29.6%, $SD = 16.2$), DE environments ended up in the bottom right corner of the graph with high DE-ratings (81.2%, $SD = 15.2$) and low UE-ratings (31.9%, $SD = 19.9$), and NM environments ended up in the bottom left corner of the graph with low DE and UE-ratings, even if slightly less sharply (respectively, 27.7%, $SD = 17.5$; 44.7%, $SD = 26.3$).

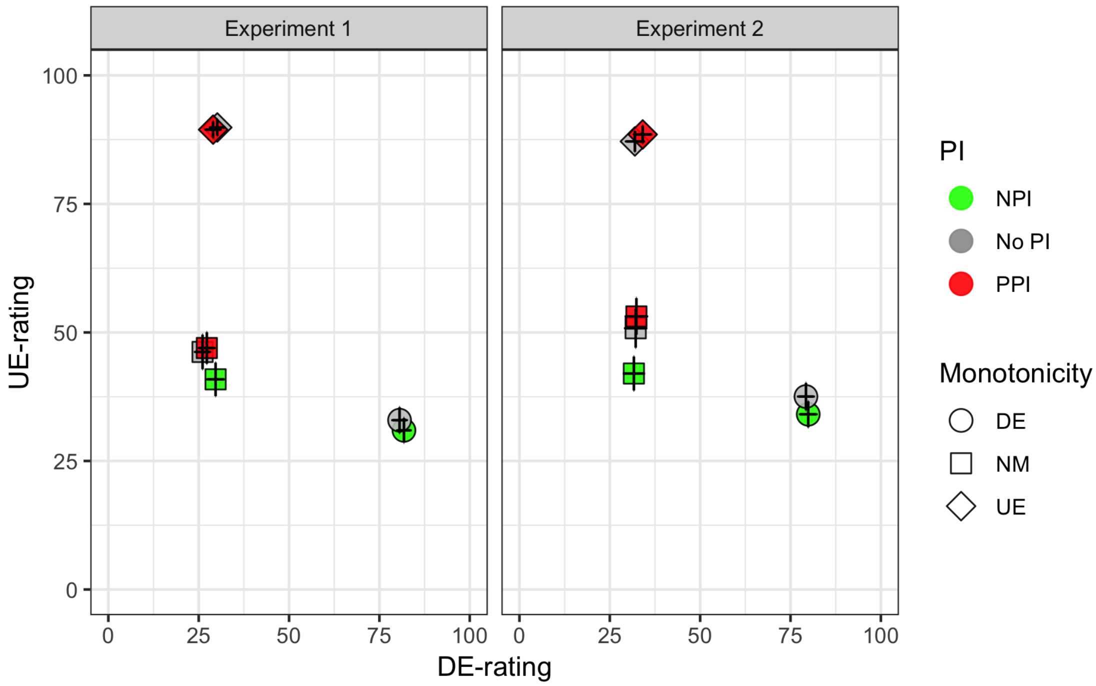


Figure 5.1: Experiments 1 and 2: Mean participants' rating of the superset to subset inference (DE-rating) and of the subset to superset inference (UE-rating) in DE, UE, and NM environments depending on whether the premise contained a PPI, an NPI, or no PI. Error bars represent standard errors.

The results are further separated depending on whether the premise contained a PPI, an NPI, or no PI, which is the core manipulation of interest: **does the presence of these items influence UE and DE ratings?** In order to answer this question, we entered ratings in a model by first transforming the ratings so that they would receive a unique directional interpretation: UE-ratings were kept untransformed, but DE-ratings were reversed (x would become $100\% - x$). These transformed measure aligns the ratings across conditions in the following sense: it measures to what extent UE inferences follows, and to what extent DE inferences do not follow. We will refer to these as directional ratings.

Focusing on NM environments, mean participants' directional ratings seem to be, numerically, influenced by the polarity items, and in particular NPIs: while the average aligned ratings without PI and with a PPI are similar (60%, $SD = 11.7$ and 59.7%, $SD = 12.2$, respectively), the presence of the NPI gave rise to lower directional ratings (55.3%, $SD = 11.2$). The pertinence of these observations were confirmed by the following (planned) analyses:

The role of NPIs in NM environments The data was subsetted to items with either an NPI or no PI in the premise that is headed by a NM quantifier (Exactly 12, Only 12). A linear mixed model was fitted on this data set's directional ratings with the PI (present or absent), Quantifier (Exactly 12, Only 12) and Inference direction (set/subset vs. subset/set) as fixed effects, and the maximal random effect structure for which convergence was achieved.[1] A comparison of this model with a reduced model without the PI as fixed effect revealed a significant effect of PI on the directional rating ($\chi^2(1) = 18.6, p < .001$). In other words, the presence of an NPI in the premise makes the participants perceive an environment as less upward-entailing as compared to when no PI is in the premise.

The role of PPIs in NM environments The data was subsetted to items with either a PPI or no PI in the premise that is headed by a NM quantifier (Exactly 12, Only 12). A linear mixed model was fitted on this data set's directional ratings with the PI (present or absent), Quantifier (Exactly 12, Only 12) and Inference direction (set/subset vs. subset/set) as fixed effects, and the maximal random effect structure for which convergence was achieved.[2] A comparison of this model with a reduced model without the PI as fixed effect revealed no significant effect of PI on the directional rating ($\chi^2(1) = .05, p = .83$). In other words, the presence of a PPI in the premise did not make the participants perceive an environment as more or less upward-entailing as compared to when no PI is in the premise in this experiment.

5.3.3 Summary and discussion

Based on the overall pattern of results and non-contentious cases, our inference measure appears to be a relevant measure of the participants' perceived monotonicity of the environment (see also Chemla et al., 2011), according to which DE and UE environments are perceived as such, and NM environments as intermediate. Our crucial result is that in NM environments, the presence of an NPI can lead participants to perceive NM environments as less UE-like. No analogous effect of PPIs was found.

Note that the main observed effect (i.e. the effect of NPIs on inferential judgments in NM environments) cannot be explained away as a regression to the mean (i.e. as more random responses than in the no PI condition leading to judgments overall closer to 50% in the PI condition). The reason why this alternative explanation is not viable is that the main effect of the NPI seems to be in making people perceive a NM as less upward-entailing, pushing them further away from the mean (50%) UE-rating than the baseline (no PI) condition is (cf. 5.1).

5.4 Experiment 2: Replication (with PIs also in conclusions)

The previous result is arguably small in size. Note however that by having environments with and without PIs, in the same experiment, we were likely not to get any result *at all*: participants could have easily figured out that PIs were irrelevant.

Furthermore, in Experiment 1 PIs were present only in the premises, to assess their role as a guide for 'future' inferences, but this creates an asymmetry between premise and conclusion, which may obscure the effect of the PI (for instance, as the PI was present in the premise but

[1] The maximal random effects structure for which convergence was achieved included by-participant intercepts and by-participant slopes for Inference direction.

[2] The maximal random effects structure for which convergence was achieved included by-participant and by-item intercept, as well as by participant slopes for Inference direction. Items (here and in the other models) corresponded to the VPs used in the sentence.

not in the conclusion, it might have been quite easy for participants to ignore it.) We thus ran Experiment 2, which was identical to Experiment 1, except that whenever a PI was present in the premise, it was also present in the conclusion.

Roughly put, this Experiment provides us with a replication of the previous result.

5.4.1 Method

5.4.1.1 Instructions and task

Instructions and task were identical to those in Experiment 1.

5.4.1.2 Material

Material was identical to those used in Experiment 1 except that the conclusion sentences of the inferences contained a PI whenever the premise did.

5.4.1.3 Participants

72 participants were recruited through Amazon Mechanical Turk (35 females). One participant was excluded from the analysis for reporting not being a native speaker of English and seven more for not showing a difference in perceived monotonicity of upward and downward entailing environments (same criteria as described in section 5.3.1.3). 64 participants were thus kept for the analysis (28 females).

5.4.2 Results

To preview, the results of Experiment 2 were qualitatively identical to the results of Experiment 1. As in Experiment 1, responses given in less than 1.4s (6% of the data) or more than 10s (7% of the data) were removed.

Training items were answered as expected: the clearly valid inference received an average rating of 93%, the clearly not valid one 9%, the intermediate one 62%. The right hand side of 5.1 summarizes the results from this experiment. As before, it represents on the x-axis mean participants' rating of superset to subset inferences (DE-ratings), and on the y-axis mean participants' rating of subset to superset inferences (UE-ratings), across three types of environments (UE, DE, and NM), depending on whether the premise and conclusion contained a PPI, an NPI, or no PI.

As in Experiment 1, if we first disregard the effect of PIs, the three types of environments UE, DE and NM behave distinctly, as expected. We also observe a similar pattern as before for the role of PIs in NM environments, with inferences without PI and with a PPI receiving similar directional ratings (59.2%, $SD = 14.9$ and 60.1%, $SD = 14.1$, respectively), and the presence of NPIs leading to lower ratings (54.7%, $SD = 12.5$).

The role of NPIs in NM environments Linear mixed model comparisons[3] as in Experiment 1 revealed a significant effect of the presence of the NPI (as compared to no PI at all) on monotonicity inferences ($\chi^2(1) = 19.04, p < .001$). In other words, we replicate the result from Experiment 1 that the presence of an NPI in the premise makes the participants perceive an environment as less upward-entailing as compared to when no PI is in the premise.

[3] Convergence was achieved for the random effects structure with by-participant intercepts and by-participant slopes for Inference direction.

The role of PPIs in NM environments Linear mixed model comparisons[4] as in Experiment 1 revealed no significant effect of the presence of the PPI (as compared to no PI at all) on monotonicity inferences ($\chi^2(1) = 1.4, p = .24$). In other words, as it was the case in Experiment 1, the presence of a PPI in the premise did not make the participants perceive an environment as more or less upward-entailing as compared to when no PI is in the premise in Experiment 2 either.

5.4.3 Summary and discussion

In Experiment 1, we tested whether the presence of NPIs and PPIs as compared to no PIs in the premise has an influence on the monotonicity inferences with conclusions without PIs. Experiment 2 differed from Experiment 1 only in that whenever a PI was present in the premise, it was also present in the conclusion. The results of Experiment 2 confirm those of Experiment 1: the presence of NPIs makes participants consider an environment less upward-entailing than when it contains no PIs, while no similar effect on monotonicity judgments is found for PPIs.

5.5 Interim discussion

NPIs have been found to influence monotonicity inferences in NM environments: they lead participants to report that these environments are less UE-like than when no PI is present (it seems that the most important part of the effect is in making people perceive NM environments as less UE rather than perceiving them as more DE, as it can be seen in 5.1). The effect seems to be more robust in NM environments than in DE environments (cf. 5.1), and we believe that this could be well-explained by *ceiling* effects in DE environments (in our experiments, as well as in Szabolcsi et al., 2008). Indeed, for DE environments, participants may report DE judgments as much as they possibly can even without an NPI, leaving little room for an NPI to make the judgments even more extreme.

What should we make of the asymmetry between NPIs and PPIs? As we found no effect of PPIs, we cannot conclude much about the role of monotonicity in PPI licensing. It could be that the monotonicity does not play a role in PPI licensing, as some authors have proposed (Denić, 2015; Progovac, 2000).

An alternative would be that monotonicity plays a role in PPI licensing too, but that we simply didn't have enough power to capture the effect of PPIs on monotonicity judgments, or that we are again observing ceiling effects. In fact, in both Experiment 1 and Experiment 2 we observe that people judge NM environments as more UE than DE (average endorsement of the UE inferences in NM environments is 44.7% ($SD = 26.3$) in Experiment 1, and 48.6% ($SD = 28.5$) in Experiment 2 , while average endorsement of DE inference in NM environments is 27.7% ($SD = 17.5$) in Experiment 1, and 32.1% ($SD = 22.6$) in Experiment 2.[5] This might have created a ceiling effect for PPIs but not for NPIs: as people perceive NM environments as more UE than DE,

[4]Convergence was achieved for the random effects structure with by-participant intercepts and by-participant slopes for Inference direction.

[5]Post-hoc comparison of a linear mixed model comparison fitted on raw responses on NM sentences without PIs with Inference direction as fixed effects with random by-participant intercepts and slopes with a reduced model with just the random effects structure revealed a significant effect of Inference direction both in Experiment 1 ($\chi^2(1) = 37.8, p < .001$) and in Experiment 2 ($\chi^2(1) = 126.35, p < .001$). In Experiment 2, the convergence was achieved with by-participant intercepts only.

NPIs could exercise their effect by making people perceive NM environments as less UE, while there might not be enough room for the same effect of PPIs.

5.6 Doubly Negative Environments

In Experiments 1 and 2, we have established that there is an influence of polarity items in NM environments. NM environments are characterized by two aspects: first, inferential judgments in these environments are difficult; second both positive and negative polarity items are licensed in these environments, at least to some extent. In the continuation of the paper, we will extend the inquiry to another type of environments with these same properties, albeit possibly for rather different reasons.

The starting point is that NPIs can appear in the so called *doubly-negative* environments that are globally upward-entailing, as in:

(21) Everyone who couldn't see **at all** failed to find the treasure.

Due to the combination of two downward-entailing functions (*n't* is DE and *every* is DE in its restrictor), the NPI *at all* ends up appearing in a global UE environment in ((21)). This example thus shows that global logical properties cannot (always) be responsible for NPI licensing. Therefore researchers who advocate a monotonicity-based approach to licensing are led to go local, i.e. propose that the system that checks the acceptability of a given PI in a sentence S has access to subconstituents of S, and that PIs are licensed if at least one of the subconstituents they are in has the right monotonicity properties (Gajewski, 2005; Homer, 2012). For concreteness, in a sentence like ((21)), this system can single out the VP of the relative clause and compute its monotonicity with respect to the position of the NPI *at all* (monotonicity is a property of functions; to evaluate what we loosely call the monotonicity of a constituent, one has to abstract over a position within this constituent, e.g. the position of the PI): this constituent turns out to be DE w.r.t. this position. As the licensing condition just requires that a PI be in at least one constituent which has the appropriate monotonicity w.r.t. its position, the NPI *at all* is licensed in ((21)). Note that for the syntactic approach, examples such as ((21)) are not problematic, as the monotonicity of the environments of PIs is not directly relevant: all that matters is that the NPI be in the right syntactic configuration with at least one appropriate operator.

There is, however, yet another interesting possibility for why NPIs are acceptable in environments like ((21)). It is possible that monotonicity inferences are so hard in these environments that people wrongly consider them DE to some extent. The NPIs would thus be licensed in these environments because of the subjective (wrong) perception of their monotonicity.

Doubly-negative environments (remember that they are locally DE but globally UE) have two properties in common with NM environments which make them interesting as a further test case for an effect of polarity items on various sorts on inferential judgments: (i) they can host both positive and negative polarity items (see ((22))); (ii) even if these environments are plainly upward entailing, we submit that the corresponding inferential judgments are not easy.

Furthermore, this inquiry could also be directly informative about the status of NPI licensing in these environments (are they licensed because the local environment is DE, or because the global environment is wrongly perceived as DE/not UE?). We will discuss this later on, as it will be easier to do so with the results in place.

5.7 Experiment 3: Doubly Negative is not Positive

This third experiment tested the effect of the PIs on monotonicity inferences in environments with two accumulating downward-entailing operators, such as ((22)). We will refer to these environments as DN in the continuation, for Doubly Negative environments. As mentioned above, both PPIs and NPIs are known to be licit in these environments (which means, for PPIs like *some*, that they can be interpreted with narrow scope under both operators; for instance, ((22a)) can be interpreted as ((22b))):

(22) a. Every alien who did not see some doves is hairy.
b. Every alien who did not see any doves is hairy.

5.7.1 Method

5.7.1.1 Instructions and task

Instructions and task were identical to those in Experiment 1 and 2.

5.7.1.2 Material

The stimuli were identical to those used in Experiment 1, except for the addition of two DN environments. These were presented with a PPI, an NPI, or no PI in the premise (and no PI in the conclusion). We thus obtained 2 [superset/subset vs subset/superset] x 12 [VPs] x (3 [UE] x 2 [PPI vs no PI] + 3 [DE] x 2 [NPI vs no PI] + 2 [NM] x 3 [NPI vs PI vs no PI] + 2 [DN] x 3 [NPI vs PI vs no PI]) = 576 inference pairs. An example of premise-conclusion pair for each of the two DN environments is in ((23)) and ((24)). These 576 items were split into three groups with 192 items, which satisfied the same conditions as the groups in Experiment 1. Participants were randomly administered to one of the three groups.

(23) Condition: DN-Every-not, subset → superset, (PPI/NPI)
a. Every alien who did not see (some/any) doves is hairy.
b. Every alien who did not see birds is hairy.

(24) Condition: DN-No-without, subset → superset, (PPI/NPI)
a. No alien spent a year without seeing (some/any) doves.
b. No alien spent a year without seeing birds.

5.7.1.3 Participants

112 participants were recruited through Amazon Mechanical Turk (69 females). Seven participants were excluded from the analysis for reporting not being a native speaker of English and 13 more for not showing much difference in perceived monotonicity of upward and downward entailing environments (same exclusion criteria as in Experiments 1 and 2). 92 participants were thus kept for the analysis (53 females).

5.7.2 Results

To preview, in Experiment 3 we mostly replicate the result from Experiments 1 and 2 for the NM environments. As for DN environments, no effect of the presence of NPIs on the global monotonicity inference was found, but interestingly, an effect of the PPI was found.

As in Experiments 1 and 2, responses given in less than 1.4s (4% of the data) or more than 10s

(13% of the data) were removed from the analysis. Training items were answered as expected: the clearly valid inference received an average rating of 92%, the clearly not valid one 8.3%, the intermediate one 67%.

5.2 represents on the x-axis mean participants' rating of superset to subset inference (DE-rating), and on the y-axis mean participants' rating of subset to superset inference (UE-rating), across four types of environments (UE, DE, NM, and DN), depending on whether the premise contained a PPI, an NPI, or no PI.

Disregarding whether and which PI was present in the premise, the four types of environments (UE, DE, NM and DN) are well-separated and they show up where they could have been expected. DN environments are quite interesting in this respect: as a reminder, DN environments are in fact plain UE environments. Nonetheless, they seem to behave in a more intermediate fashion, and they are much closer to NM than to UE environments.

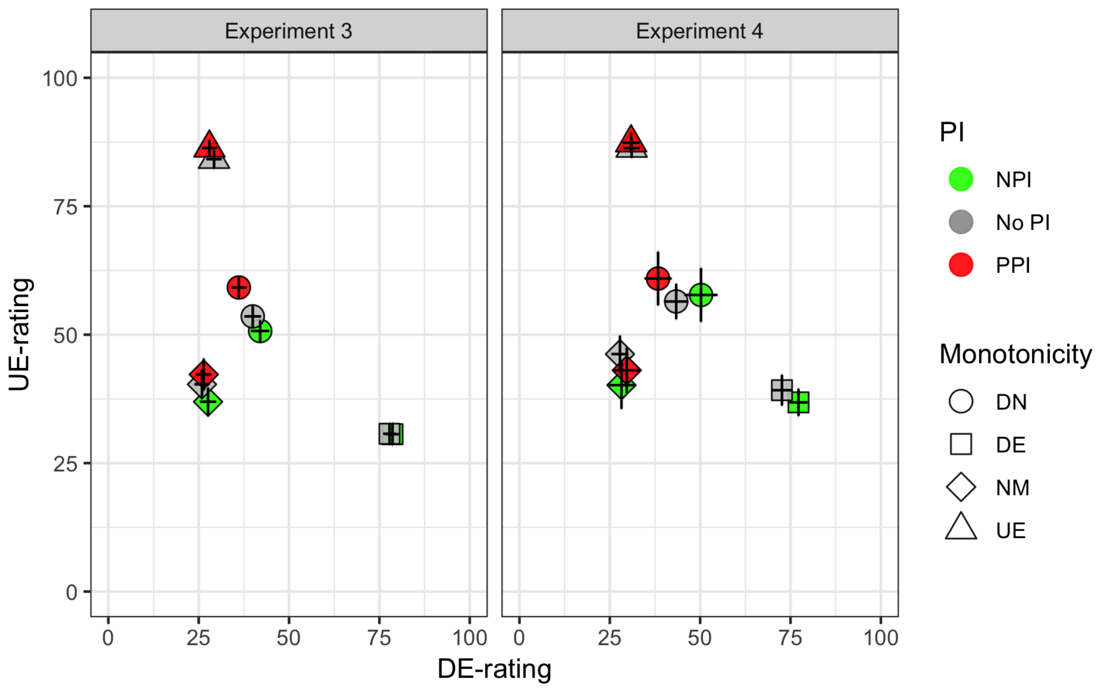


Figure 5.2: Experiment 3 and 4: Mean participants' rating of the superset to subset inference (DE-rating) and of the subset to superset inference (UE-rating) in DE, UE, NM, and DN environments depending on whether the premise contained a PPI, an NPI, or no PI. Error bars represent standard errors.

Looking first at the replication of the results from Experiments 1 and 2 in NM environments, mean participants' directional ratings were (i) 55.9%($SD = 13.4$) when NPI is in the premise, (ii) 57.9%($SD = 14$) when PPI is in the premise, and (iii) 57.1%($SD = 13.3$) when no PI is in the premise.

The role of NPIs in NM environments Linear mixed model comparisons[6] as in Experiment 1 revealed a borderline effect of the presence of the NPI (as compared to no PI at all) on

[6]Convergence was achieved for the random effects structure which included by-participant intercepts only.

monotonicity inferences ($\chi^2(1) = 2.95, p = .086$) in NM environments. In other words, even though Experiment 3 is like Experiments 1 and 2 in that numerically the NPI in the premise makes the participants perceive NM environments as less upward-entailing as compared to when no PI is in the premise, this effect doesn't reach significance in Experiment 3, unlike in Experiments 1 and 2.

The role of PPIs in NM environments Linear mixed model comparisons[7] as in Experiments 1 and 2 revealed no significant effect of the presence of the PPI (as compared to no PI at all) on monotonicity inferences ($\chi^2(1) = 1.9, p = .17$) in NM environments. In other words, as it was the case in Experiments 1 and 2, the presence of a PPI in the premise did not make the participants perceive non-monotonic environment as more or less upward-entailing as compared to when no PI is in the premise.

Moving to DN environments, mean participants' directional ratings were (i) 53.7%(SD = 16.8) when NPI is in the premise, (ii) 61.8%(SD = 14.4) when PPI is in the premise in, and (iii) 56.9%(SD = 14.7) when no PI is in the premise.

The role of NPIs in DN environments Linear mixed model comparisons as for the NM environments was done for DN environments[8] and it revealed no significant effect of the presence of the NPI (as compared to no PI at all) on monotonicity inferences ($\chi^2(1) = 1.96, p = .16$). In other words, we find no effect of NPI on monotonicity inferences in DN environments.

The role of PPIs in DN environments Linear mixed model comparisons[9] revealed a significant effect of the presence of the PPI (as compared to no PI at all) on monotonicity inferences ($\chi^2(1) = 11.8, p < .001$). In other words, even though the presence of PPIs had no effect on monotonicity inferences in NM environments in Experiments 1-3, it did influence monotonicity inferences in DN environments.

5.7.3 Summary and discussion

In Experiment 3, in addition to replicating the results from Experiments 1 and 2 (although in one case, only numerically), we tested whether the presence of NPIs and PPIs as compared to no PIs in DN environment has an influence on global monotonicity inferences. We found no significant effect of NPIs, but an effect of PPIs on the global monotonicity inferences was found in DN environments (which are upward entailing). Why do we see the effect PPIs on inferential judgments in DN but not in NM environments?

This could be related to the fact that the baseline inferential judgments may leave more room for PPI effects in DN environments than in NM environments. Let us explain. While neither of the two environments are DE, DN environments are judged on average as more DE

[7]Convergence was achieved for the random effects structure which included by-participant intercepts and by-participant slopes for Inference direction.

[8]Convergence was achieved for the random effects structure which included by-participant and by-item intercepts and by-participant slopes for Quantifier and Inference direction.

[9]Convergence was achieved for the random effects structure which included by-participant intercepts and by-participant slopes for Quantifier and Inference Direction.

than NM environments (average DE-rating in DN is 40.1%, with *SD* = 23.7, and average DE-rating in NM environments is 28.5%, with *SD* = 20.4). This means that there is more room for improvement in terms of DE-rating in DN environments as opposed to NM environments. Accordingly, a PPI effect pushing judgments away from this mistake may be easier to observe for DN environments.

However, this result must be interpreted with care, as there is an alternative explanations for why we see the effect of PPIs in DN but not in NM environments. This alternative explanation comes from the fact that PPIs like *some* can take an exceptional wide scope. For instance, ((25a)) has an interpretation according to which *some* takes the widest scope in the sentence (this interpretation is paraphrased in ((26))). Note that under this interpretation, the subset to superset inference from ((25a)) to ((25b)) still follows.

(25) a. Every alien who did not see some doves is hairy.
b. Every alien who did not see birds is hairy.

(26) Some doves are such that every alien who didn't see them is hairy.

This means that, if for whichever reason the wide scope of *some* is easier to obtain in DN environments than in NM environments, and if the wide scope interpretation of *some* makes it easier to see that the inference from ((25a)) to ((25b)) follow, one could explain the effect of PPI in DN environments without relating PPI licensing to monotonicity inferences.

Let us now discuss another puzzling aspect of the results from Experiment 3, namely, that we don't see an effect of NPIs in DN environments. Why would this be?

The first, uninteresting, possibility, is that we simply didn't have enough power to capture it. As DN environments are arguably the most complex among the environments tested in the experiments, they might have lead to a larger within-participant variation (this seems to clearly be the case in Experiment 4, even though less so in Experiment 3), requiring more power to capture the effect if it was there.

There is, however, yet another possibility, compatible with theories according to which an NPI is licensed if there is at least one constituent which has the right downward monotonicity properties. If these theories are on the right track, the NPI in DN environments can in principle have conflicting influences, depending on the constituent in which the NPI is thought to be licensed: (i) it might influence the perception of the local environment as more DE, thus leading to the overall better perception of the global environment as more UE than in the no PI condition, or (ii) it might influence the perception of the global environment as more DE than in the no PI condition. If both are happening to more or less the same extent, it is in fact not surprising not to find a significant effect of the NPI in DN environments.

5.8 Experiment 4: Doubly Negatives, a replication (with a different arrangement of the items)

5.8.1 Method

5.8.1.1 Instructions and task

Instructions and task were identical to those in Experiments 1, 2 and 3.

5.8.1.2 Material

Materials were the same as those used in Experiment 3, with two differences. First, the total number of items was reduced to 480 by reducing the number of different VPs from 12 to 10. Second, the participants were split into four groups (instead of three) in such a way that each participant sees a given environment either with an NPI, or with a PPI, or without a PI. Because of this, there were 2 [superset/subset vs subset/superset] x 10 [VPs] x (3 [UE] + 3 [DE] + 2 [NM] + 2 [DN]) = 200 items per group.

5.8.1.3 Participants

81 participants were recruited through Amazon Mechanical Turk (43 females). Four participants were excluded from the analysis for reporting not being a native speaker of English and six more for not showing much difference in perceived monotonicity of upward and downward entailing environments (same exclusion criteria as in Experiments 1-3). 71 participants were thus kept for the analysis (36 females).

5.8.2 Results

As in Experiments 1-3, we removed responses given in less than 1.4s (8% of the data) or more than 10s (10% of the data). Training items were answered as expected: the clearly valid inference received an average rating of 89.5%, the clearly not valid one 8.2%, the intermediate one 60.7%.

5.2 represents on the x-axis mean participants' rating of superset to subset inference (DE-rating), and on the y-axis mean participants' rating of subset to superset inference (UE-rating), across four types of environments (UE, DE, NM, and DN), depending on whether the premise contained a PPI, an NPI, or no PI.

Disregarding for the time being the role of the PIs on inferential judgments, the four environments (DE, UE, NM, DN) are well separated: these results seem to be qualitatively identical to the results in Experiment 3.

Looking again first at the effect of PIs in NM environments, mean participants' directional ratings are (i) 56.3%($SD = 11.6$) when NPI is in the premise, (ii) 56.4%($SD = 11.1$) when PPI is in the premise in, and (iii) 59.3%($SD = 15.3$) when no PI is in the premise.

The role of NPIs in NM environments Linear mixed model comparisons[10] as in Experiments 1-3 revealed a significant effect of the presence of the NPI (as compared to no PI at all) on monotonicity inferences ($\chi^2(1) = 7.8, p < .01$) in NM environments. In other words, we replicate the results from Experiments 1 and 2 that the presence of an NPI in the premise makes the participants perceive NM environments as less upward-entailing as compared to when no PI is in the premise.

Furthermore, with the results of the four experiments in place, we conducted a post-hoc meta-analysis over the four Experiments. Linear mixed model comparisons[11] as above revealed a significant effect of the presence of the NPI (as compared to no PI at all) on monotonicity inferences in NM environments when pooling the results of all experiments ($\chi^2(1) = 38.4, p < .001$).

[10] Convergence was achieved for the random effects structure which included by-participant and by-item intercepts as well as by-participant slopes for Inference direction.

[11] Convergence was achieved for the random effects structure which included by-participant and by-item intercepts as well as by-participant slopes for Inference direction and for PI.

The role of PPIs in NM environments Linear mixed model comparisons[12] as in Experiments 1-3 revealed no significant effect of the presence of the PPI (as compared to no PI at all) on monotonicity inferences ($\chi^2(1) = 2.6, p = .10$) in NM environments. In other words, as it was the case in Experiments 1-3, the presence of a PPI in the premise did not make the participants perceive NM environments as more or less upward-entailing as compared to when no PI is in the premise.

Again, with the results of the four experiments in place, we conducted a post-hoc meta-analysis over the four experiments. Linear mixed model comparisons[13] revealed no significant effect of the presence of the PPI (as compared to no PI at all) on monotonicity inferences in NM environments across four experiments ($\chi^2(1) = 0.8, p = .38$).

In DN environments, mean participants' directional ratings were (i) 52.6%($SD = 25.2$) when NPI is in the premise, (ii) 61.6%($SD = 17.4$) when PPI is in the premise in, and (iii) 56.3%($SD = 22.4$) when no PI is in the premise.

The role of NPIs in DN environments Linear mixed model comparisons as for the NM environments was done for DN environments. This time, the result was ambiguous, as convergence was achieved for two equally complex random effects structures. According to the first one[14], the presence of the NPI (as compared to no PI at all) has a significant effect on monotonicity inferences ($\chi^2(1) = 27.9, p < .001$). According to the second one[15], it does not ($\chi^2(1) = 2.62, p = .11$).

We conducted a post-hoc meta-analysis to evaluate this effect pooling Experiments 3 and 4, in which these environments were tested. Linear mixed model comparisons[16] revealed a significant effect of the presence of the NPI (as compared to no PI at all) on monotonicity inferences in DN environments ($\chi^2(1) = 7.2, p < .01$).

The role of PPIs in DN environments Linear mixed model comparisons lead to an ambiguous result, as convergence was achieved for two equally complex random effects structures. According to the first one[17], the presence of the PPI (as compared to no PI at all) has a significant effect on monotonicity inferences ($\chi^2(1) = 22.3, p < .001$). According to the second one[18], it does not ($\chi^2(1) = 2.1, p = .15$).

We conducted a post-hoc meta-analysis to evaluate this effect pooling Experiments 3 and

[12] The maximal random effects structure for which convergence was achieved had by-participant intercepts as well as by-participant slopes for Inference direction.

[13] Convergence was achieved for the random effects structure which included by-participant intercepts as well as by-participant slopes for Inference direction.

[14] In the first model, convergence was achieved with the random effects structure which included by-participant intercepts and by-participant slopes for Inference direction.

[15] In the second model, convergence was achieved with the random effects structure which included by-participant intercepts and by-participant slopes for Quantifier.

[16] Convergence was achieved for the random effects structure which included by-participant intercepts as well as by-participant slopes for Inference direction and PI.

[17] In the first model, convergence was achieved with the random effects structure which included by-participant intercepts and by-participant slopes for Inference direction.

[18] In the second model, convergence was achieved with the random effects structure which included by-participant intercepts and by-participant slopes for PI.

4, in which these environments were tested. Linear mixed model comparisons[19] revealed a significant effect of the presence of the PPI (as compared to no PI at all) on monotonicity inferences in DN environments ($\chi^2(1) = 10.7, p < .01$).

5.8.3 Summary and discussion

In Experiment 4, we replicate the effect of the presence of the NPI on monotonicity inferences in NM environments. This effect is further supported by the meta-analysis performed on the results of the four experiments reported in this paper. As in the previous experiments, in Experiment 4 we find no effect of PPI on monotonicity inferences in NM environments (and a meta-analysis does not say differently).

As for DN environments, the results of Experiment 4 provide some evidence for the effect of both NPIs and PPIs on inferences, which evidence is further confirmed by the joint analysis of Experiment 3 and 4, suggesting that the NPI effect was missed in Experiment 3 for power reasons. However, it is important to remember that this effect currently enjoys lower replicability than other effects reported across these experiments.

5.9 Conclusion and discussion

In the four experiments reported in this paper, it was found that polarity items affect reasoning judgments. These PI manipulations are subtle on the surface, and give rise to accordingly small and subtle effects on upward and downward monotonicity inferences. These effects were not found in previous investigations with simpler cases (plain DE and UE environments, see Szabolcsi et al., 2008), and likewise, we found that the effects are smaller or absent in those environments. However, in more complex cases where inferential judgments are more difficult, in particular in NM environments, polarity items have been found to influence inferential (monotonicity) judgments across multiple experiments. These results thus reveal the influence on reasoning tasks of subtle and apparently minor choices of closed class words.

Interestingly, the current results also document the fact that monotonicity inferences may be rather difficult to assess in inferential judgments tasks (Geurts and van Der Slik, 2005). Whether this is simply a consequence of performing the experimental task in question is open for discussion, but surely it raises challenges for the question of how people derive and understand the truth-conditions of sentences for efficient communication, noting that as soon as one understands the truth-conditional meaning of a sentence, one should be able to see what is entailed by it. We note here that polarity items can help filter out some misunderstandings, if they can help out signaling monotonicity properties.

A final, more technical issue one may raise is how our results bear on the question of why NPIs are licensed in DN environments. There is, in fact, a surprising result which might speak to this question. We found that the environments in which NPIs are thought to be licensed are not all perceived as downward entailing, but all of them are perceived as 'non upward entailing'. This is perhaps unsurprising for DE and NM environments, but it is remarkable for DN environments, which are in fact upward entailing. Furthermore, some results suggest that the presence of an NPI makes one perceive a DN environment as more downward-entailing (which, again, it is not). This opens the possibility that the acceptability of NPIs in such environments is not

[19] Convergence was achieved for the random effects structure which included by-participant and by-item intercepts as well as by-participant slopes for Inference direction and PI.

(only) due to a syntactic relationship with a negative licensor or to the downward monotonicity of the local environment, but rather (at least in part) to the perception that the environment is, at a global level, not upward entailing. This is a significant departure from current approaches (although see Chemla et al., 2011). This perspective emerges here from the systematic collection of inferential and acceptability judgments, and it could naturally be put to further tests (although testing more environments in this way is more costly than through traditional introspective judgments). This theoretical option may also illustrate a particular type of cognitive approach to linguistic generalizations in general, in which subjective and potentially 'fallacious' judgments have their say in grammar and grammatical theorizing.

Acknowledgement

We wish to thank Lewis Bott, Alexandre Cremers, Luka Crnič, Danny Fox, Philippe Schlenker, Benjamin Spector, Anna Szabolcsi, Lyn Tieu for comments and discussions. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n.313610 and was supported by ANR-17-EURE-0017 and by a travel grant from the ESF as part of the Euro-Xprag network.

Chapter 6

Intervention effects in NPI licensing

This chapter is a reproduction of the following article:

Denić, M., Chemla, E., & Tieu, L. (2018). Intervention effects in NPI licensing: A quantitative assessment of the scalar implicature explanation. *Glossa: A Journal of General Linguistics, 3(1), 49.*

Abstract

This paper reports on five experiments investigating intervention effects in negative polarity item (NPI) licensing. Such intervention effects involve the unexpected ungrammaticality of sentences that contain an *intervener*, such as a universal quantifier, in between the NPI and its licensor. For example, the licensing of the NPI *any* in the sentence **Monkey didn't give every lion any chocolate* is disrupted by intervention. Interveners also happen to be items that trigger scalar implicatures in environments in which NPIs are licensed (Chierchia, 2004, 2013). A natural hypothesis, initially proposed in Chierchia (2004), is that there is a link between the two phenomena. In this paper, we investigate whether intervention effects arise when scalar implicatures are derived.

6.1 Negative polarity items and intervention effects

Negative polarity items (NPIs) are expressions that are sensitive to the logical properties of the environment in which they occur. Examples of NPIs in English include *any*, *anybody*, *anywhere*, and *ever*. A generalization that successfully captures the distribution of NPIs is that these items are acceptable in downward-entailing (DE) environments (Fauconnier, 1975a; Ladusaw, 1979), i.e. environments that license inferences from sets to subsets. For example, (1a) entails (1b), with *chocolate muffins* denoting a subset of *muffins*, and the NPI *any* is acceptable. Conversely, the environment in (2) is not DE ((2a) does not entail (2b)), and the NPI *any* is not licensed.

(1) a. Ana didn't bake any muffins today.
 b. Ana didn't bake any chocolate muffins today.

(2) a. Ana baked (*any) muffins today.
 b. Ana baked (*any) chocolate muffins today.

Expressions that create a DE context for the NPI, such as negation, are called NPI licensors. Just as with negation in (1), we can see that *without*, in (3), and the restrictor of the universal quantifier, in (4), are also NPI licensors.

(3) a. John came to the party without any muffins.
b. John came to the party without any chocolate muffins.

(4) a. Everyone who tried any muffins that Ana baked loved them.
b. Everyone who tried any chocolate muffins that Ana baked loved them.

Certain DE environments, on the other hand, resist this generalization and do not license NPIs. This happens systematically when certain elements, so-called *interveners*, occur in between the licensor and the NPI, giving rise to so-called intervention effects (Linebarger, 1987; Krifka, 1995; Chierchia, 2004; Beck, 2006; Guerzoni, 2006). For example, the universally quantified NP is an intervener, but the definite NP is not: consider the two pairs in (5) and (6), which differ only in the intervener status of the indirect object. (5a) and (6a) entail (5b) and (6b), respectively, yet adding the NPI in (5a,b) leads to degradation, while adding the NPI in (6a,b) does not.

(5) a. Monkey didn't give every rabbit (*any) juice.
b. Monkey didn't give every rabbit (*any) strawberry juice.

(6) a. Monkey didn't give the rabbits any juice.
b. Monkey didn't give the rabbits any strawberry juice.

Another intervener is the conjunction *and*: when conjunction is in the scope of negation, (7a) entails (7b), yet the NPI *any* is not licensed.

(7) a. Ana didn't bake both cookies and (*any) muffins.
b. Ana didn't bake both cookies and (*any) chocolate muffins.

Strikingly, interveners form a natural class: they are items that trigger so-called scalar implicatures in DE environments (Chierchia, 2004). Roughly, scalar implicatures are optional inferences arising when (i) a sentence can be argued to have a minimally different *alternative*, obtained for instance by replacing some lexical item with a similar one; for instance, *some* and *all*, or *or* and *and* would count as similar, or *scale-mates* (see Horn, 1972; Gazdar, 1979, as well as Katzir, 2007 and Fox and Katzir, 2011 for recent discussions about the derivation of alternatives), and (ii) the sentence is consistent with the alternative being false (see Grice, 1975; Sauerland, 2004; van Rooij and Schulz, 2004; Schulz and van Rooij, 2006; Spector, 2006, 2007; Chierchia et al., 2012; Franke, 2011; Bergen et al., 2016 for refinements and discussions from a variety of perspectives). In these conditions, the negation of the alternative may be added to the meaning of the sentence, as a scalar implicature. This is best understood through an example. (8a) can be argued to have (8b) as an alternative, obtained by replacing the universal quantifier *every* with the existential quantifier *any*. Because the environment is DE, the alternative (8b) is logically stronger than the sentence (8a), and therefore the negation of the alternative (8c) is compatible with the sentence, hence it may be added as a scalar implicature of the sentence.

(8) a. Sentence: Monkey didn't give every rabbit juice.
b. Alternative: Monkey didn't give any rabbit juice.
c. Scalar implicature: Monkey gave some of the rabbits juice.

Given that interveners are items that trigger scalar implicatures in DE environments, a natural hypothesis is that there is a link between the two. To make things slightly more concrete, one could say that scalar implicatures add a non-DE component to the meaning of a sentence, and for precisely this reason they may disrupt the licensing of an NPI.[1] The possible connection between implicatures and NPI licensing has been pursued by Chierchia (2004, 2013) in detail (for proposals that do not relate intervention effects to implicatures, see for example Beck, 2006 and Guerzoni, 2006). The proposal that implicatures are the cause of intervention effects is appealing in terms of its explanatory power. In particular, scalar implicatures are not the only inferences reported to give rise to intervention effects; Homer (2008) argues that presuppositions also give rise to intervention effects for the same reason, namely that they can disrupt the downward-entailingness of the licensing environment.

In this paper, we will focus on the proposal that implicatures might be the cause of intervention effects in sentences like (5). The question is whether such intervention effects arise when scalar implicatures are derived. An immediate reason to doubt that this is the case is that implicatures tend to be volatile, while intervention effects are reported to create categorically bad sentences. To explain the purportedly categorical ungrammaticality of intervention sentences, Chierchia (2004) proposes that the strong meaning of a sentence, i.e. its meaning enriched with an implicature, is available automatically in parallel with its plain implicature-less meaning, and crucially, it is this *strong* meaning against which NPI licensing must be checked.[2]

We will start by assessing whether the initial objection to the implicature-based theory is empirically justified to begin with. We will develop appropriate methods to measure intervention effects and assess what potential volatility they themselves may exhibit (Experiment 1). We will observe that they do not lead to categorical ungrammaticality as usually assumed, which reduces the initial challenge to the implicature theory, as it actually connects the two volatile phenomena.

The results of Experiment 1 are thus a motivation to abstract away from the specific details of Chierchia's proposal regarding the exact licensing mechanism of NPIs and how it interacts with implicature derivation. Instead, we will consider the expectations that arise from a pos-

[1] More precisely, the relevant implicatures in typical intervention configurations may turn a DE environment into an overall non-monotonic environment, i.e. an environment that does not license inferences from sets to supersets or subsets. Example (9) involves a non-monotonic environment (albeit not by way of implicatures); it is known that NPIs in such environments are not perfectly acceptable (see Rothschild, 2006 and Crnič, 2014 for discussion, and Chemla et al., 2011 for quantitative data).

(9) a. Exactly two boys baked muffins. $\not\Rightarrow$ Exactly two boys baked chocolate muffins.
    b. Exactly two boys baked muffins. $\not\Rightarrow$ Exactly two boys baked.

[2] In light of this issue of the automaticity of strong (implicature-enriched) meanings, an anonymous reviewer asks whether processing data on implicatures may be relevant. In fact, the question of whether implicatures are derived automatically or not is orthogonal to the question of whether variability in implicature derivation can explain variability in intervention effects. The question only becomes relevant if (i) implicatures are derived automatically *and* (ii) one stipulates that NPI licensing must be checked before any implicature cancellation can occur. If this were the case, however, intervention effects would be expected to be categorical, and not volatile. As we will see, the results of our Experiment 1 run counter to this expectation. Regarding the processing of implicatures, we will merely note here that the majority of previous experimental work suggests that interpreting a sentence without its implicatures is equivalent to not deriving the implicatures rather than cancelling them; for instance, Bott and Noveck (2004) and Bott et al. (2012) provide evidence that participants who give implicature-less responses are faster to respond than those who give implicature-based responses, and Cremers and Chemla (2014) report similar results for 'indirect' scalar implicatures, which are the kind that concern us here (though see also Romoli and Schwarz 2015).

sibly more general family of theories according to which implicatures are the cause of intervention effects, hereafter referred to as the scalar implicature (SI) theory of intervention effects. In particular, we will propose to evaluate whether the variability of one can be traced to the variability of the other.

To do so, we will make use of a range of common paradigms to compare spontaneous implicature derivation and grammaticality judgments of intervention sentences within individual participants. To preview, a rather mixed empirical picture will emerge across the different experiments. Experiments 2 and 3, using a picture selection task and a covered picture task to measure implicature derivation, reveal no correlation between individual implicature derivation rates and sensitivity to intervention effects. Experiment 4, employing a training paradigm to train participants to derive or not to derive implicatures, likewise reveals no correlation between implicatures and intervention. Finally, Experiment 5, which tests participants *simultaneously* on implicatures and intervention by observing their repair strategies for intervention sentences, reveals evidence that people may react to intervention sentences by effectively shutting off implicatures. As we will discuss in Section 6.7, the mixed empirical landscape revealed by the full set of experiments raises new challenges for theories of intervention.

The data and R analysis scripts (R Core Team, 2016) for the experiments are available online at `https://tinyurl.com/y6m68wdq`.

6.2 Experiment 1: Grammaticality judgments of intervention effects

As was pointed out in Section 1, the SI theory of intervention effects is challenged by the fact that scalar implicatures can be suspended, while sentences with the NPI in an intervention configuration are reported to be categorically ungrammatical. The goal of Experiment 1, which used an acceptability judgment task, was two-fold: confirm experimentally the presence of intervention effects in NPI licensing, and assess whether intervention configurations indeed lead to categorical ungrammaticality, as is usually assumed.

6.2.1 Participants

54 participants (15 female) were recruited on Amazon Mechanical Turk, and were paid $1.80 for their participation. One participant was excluded from analysis for not being a native speaker of English.

6.2.2 Procedure and materials

Participants were directed to a web-based acceptability judgment task, hosted on Alex Drummond's Ibex platform for psycholinguistic experiments. Participants were told that they would read sentences (about animal characters) produced by Zap, an alien learning English, and that they were to judge how these sentences sounded on a scale from 1 to 5, with 1 being completely odd and 5 being completely okay. Participants registered their ratings by clicking on the appropriately numbered box (see Figure 6.1 for an example of the response buttons).

Participants first saw two practice trials, one involving a clearly well-formed sentence and one involving a clearly ill-formed sentence, accompanied by suggested ratings of 5 and 1, respectively. The purpose of these examples was to demonstrate to the participants that Zap was indeed capable of producing both acceptable and odd sentences. Participants then began the test phase of the experiment, the first two items of which were identical to the two practice trials. These were then followed by the 48 test trials schematically described in Table 6.1. We

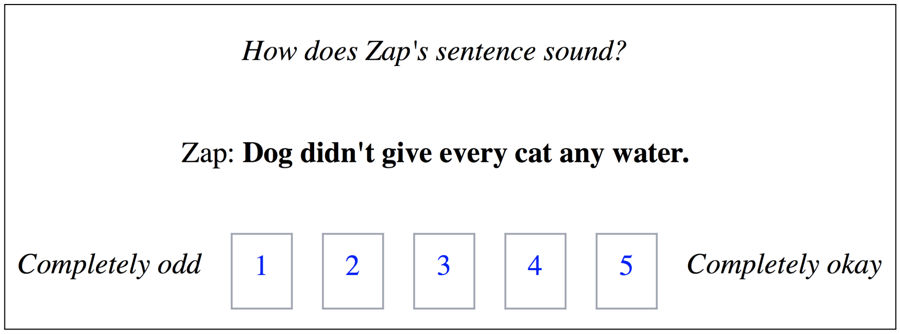


Figure 6.1: Experiment 1: An example of a test item with the NPI *any* in an intervention configuration

manipulated two factors, crossing NPI (present vs. absent) with the six item types listed in Table 6.1. The 48 trials were presented in randomized order.

Table 6.1: Summary of trial types

Item type	Linguistic environment	Example sentence	Repetitions
Target	Negation, +intervener	Monkey didn't give every bear (any) pie.	16 [+NPI], 4 [-NPI]
Downward-entailing (DE)	Negation, -intervener	Monkey didn't give the rabbits (any) juice.	4 [+NPI], 4 [-NPI]
Non-monotonic (NM)	Scope of *exactly two*	Exactly two elephants ate (any) cake.	4 [+NPI], 1 [-NPI]
Upward-entailing (UE) - Simple	Positive sentence	Lion drank (any) coffee.	4 [+NPI], 1 [-NPI]
Upward-entailing (UE) - Complex	Nuclear scope of *every*, -negation in restrictor	Every rabbit who was hungry ate (any) chocolate.	4 [+NPI], 1 [-NPI]
Upward-entailing (UE) - Illusory	Nuclear scope of *every*, +negation in restrictor	Every bear who didn't have ice cream ate (any) pie.	4 [+NPI], 1 [-NPI]

As seen in Table 6.1, in addition to the target sentences, participants saw sentences involving DE and non-monotonic environments, as well as sentences involving different kinds of upward-entailing (UE) environments, which license inferences from sets to supersets. The UE sentences were not expected to license the NPI, and included simple UE sentences (UE-Simple), more complex UE sentences (UE-Complex), and UE sentences with a DE operator but in an irrelevant position (UE-Illusory). The particular number of repetitions of each of the non-target trial types was chosen in such a way as to balance the overall number of degraded sentences (i.e. UE-simple, UE-complex, UE-illusory, NM[+NPI]) and non-degraded sentences (i.e. DE[+NPI], [-NPI]).

The target sentences involved negative ditransitives built on the structural template '<Subject NP> didn't give every <Indirect Object NP> <Direct Object NP>'. For the subject and indirect object NPs, we randomly chose animals from the following list: lion, rabbit, cat, dog, giraffe, bear, monkey, elephant. The direct object NPs were randomly chosen from the following list: juice, tea, milk, ice cream, pie, cheese, cake, honey. We opted to use mass nouns, in order to avoid any confound related to the plurality of objects.

6.2.3 Results

6.2.3.1 Exclusions

We excluded participants whose mean judgment for uncontroversially good cases (DE [+NPI] and all [-NPI] items) was *lower* than their mean judgment for uncontroversially bad cases (all UE [+NPI] items). This allowed us to ensure that participants were doing the task appropriately and understood the way the scale was supposed to be used. This criterion led to the exclusion of one participant, whose mean judgment on uncontroversially bad cases was 3.25, compared to a mean judgment of 2.69 for the uncontroversially good cases. As for the remaining participants, the mean judgment for the uncontroversially bad cases was 2.07, while the mean judgment for the uncontroversially good cases was 4.47.

6.2.3.2 Targets

Figure 6.2 presents the mean judgments across different environments from the remaining 52 participants (collapsing the three UE conditions); Figure 6.3 presents the same data in more detail for the critical sentences with an NPI. We are mostly interested in comparing the acceptability of the NPIs, in different environments. To do so, we computed a measure of this acceptability that factors out possible effects coming from the environment itself, independently of the NPIs. Specifically, for each participant and each environment, we calculated the log-ratio of the mean acceptability ratings for the [+NPI] targets over their [-NPI] counterparts, schematically: $\log \frac{\text{mean([+NPI])}}{\text{mean([-NPI])}}$. These log-ratios measure the degradation *due to the NPI*, for each participant in each environment; a comparison of log ratios for two environments (below, through paired two-tailed t-tests) thus determines in which environment the introduction of an NPI yielded a higher degradation, i.e. which condition was judged worse.

First, we observe, as we expect, that NPIs are rated better in DE environments than in NM environments ($t(51) = 9.7, p < 0.001$), and that they are more acceptable in NM environments than in each of the UE environments (UE-simple: $t(51) = 4.41, p < 0.001$, UE-complex: $t(51) = 3.61, p < 0.001$, UE-illusory: $t(51) = 2.45, p = 0.02$). This is as it should be, given the known intermediate status of NPIs in NM environments (Rothschild, 2006; Chemla et al., 2011; Crnič, 2014).

Second, intervention effects were detected, in the sense that the critical target sentences were judged worse than the DE sentences ($t(51) = 6.8, p < 0.001$). Crucially, the degradation was not as strong as for plain violations in UE environments, with sentences in the target environments judged better than sentences in each of the UE [+NPI] conditions (UE-Simple ($t(51) = -7.9, p < 0.001$); UE-Complex ($t(51) = -6.5, p < 0.001$); UE-Illusory ($t(51) = -5.45, p < 0.001$)). Finally, the critical target sentences were also judged better than the NM sentences ($t(51) =$

$-3.39, p = 0.001$).[3,4]

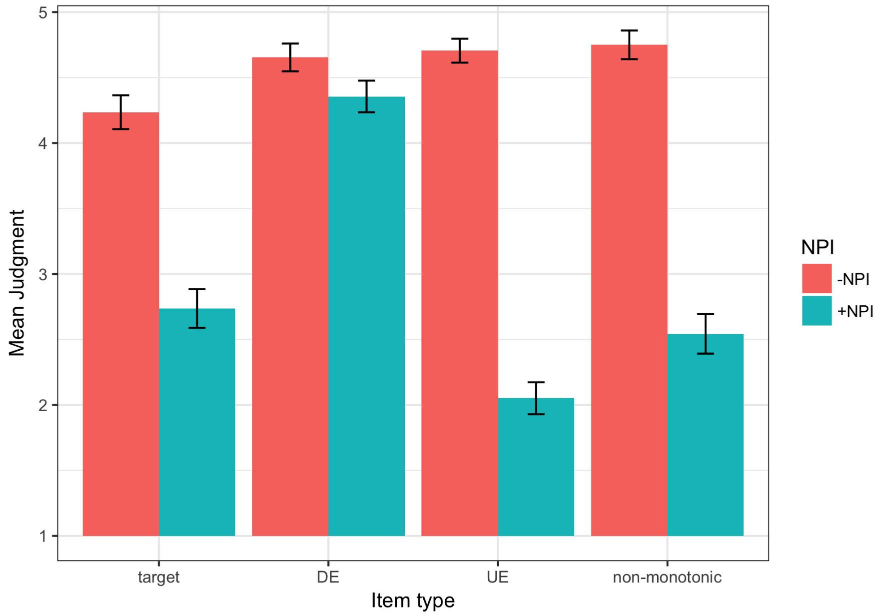


Figure 6.2: Experiment 1: Average response for each sentence type: target, downward-entailing, combined upward-entailing, and non-monotonic sentence types

6.2.4 Discussion

Intervention effects were found to be NPI licensing violations, as the NPIs in the critical target sentences were judged as less acceptable than those in the DE sentences. But the violations that they yielded were weaker than other licensing violations, as they were still judged better than NPIs in the plain UE or NM sentences. This previously unnoticed intermediate status of intervention sentences in fact eliminates the initial objection to implicature theory: that implicatures are volatile while intervention effects are not. This judgment pattern of intervention effects is thus interestingly compatible with the SI theory; intervention configurations may create NPI violations (say, because they create NM environments, see footnote 1), but this violation should only occur when the implicature is derived. Testing the dependence of the violation on the presence of the implicature is the goal of the following experiments.

[3]An anonymous reviewer asks whether the target sentences might have been rated better than the NM sentences due to a sort of satiation effect, as participants saw more targets than they did NM items (cf. Table 6.1). In a post-hoc analysis, we observed that even when we restricted the analysis of the critical targets to the very first four occurrences, all of the significant differences remained.

[4]The reported p-values are reported unadjusted, so that the reader can check that the reported differences remain significant at various thresholds and with the application of Holm-Bonferroni corrections for multiple comparisons.

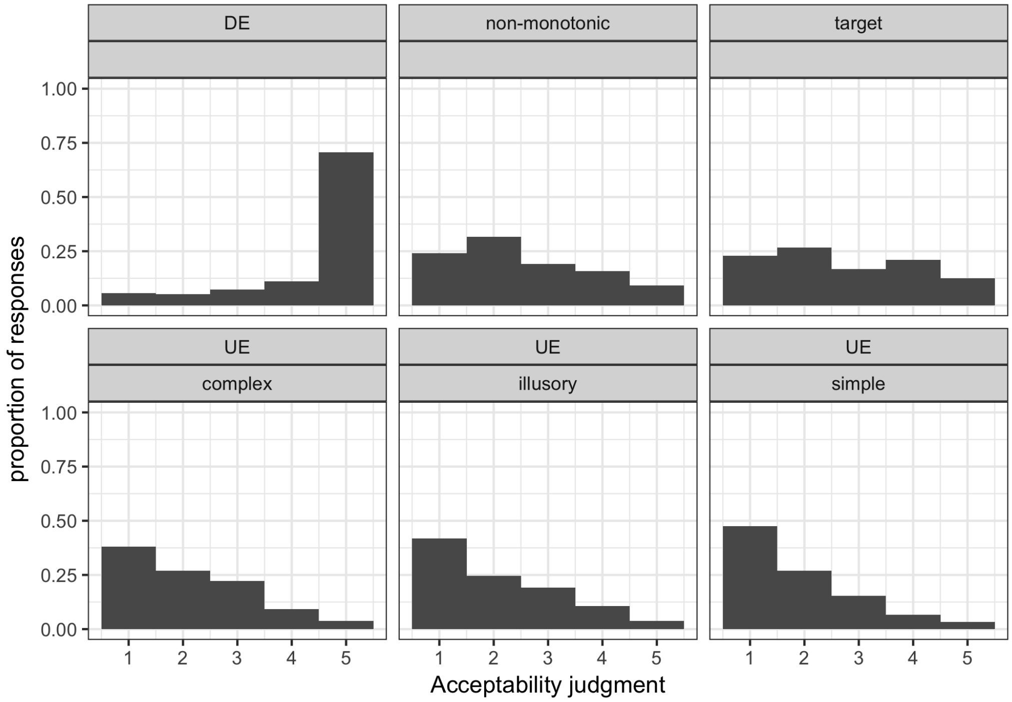


Figure 6.3: Experiment 1: Distribution of responses (1-5) for each sentence type in the [+NPI] condition

6.3 Experiment 2

In the first experiment it was shown that sentences containing NPIs in intervention configurations are judged better than sentences with NPIs in contexts that uncontroversially fail to license NPIs. This result goes well with the proposal that intervention effects are caused by scalar implicatures, but the experiment did not directly investigate the link between the two phenomena (because implicatures were not tested at all). Our next four experiments set out to investigate the presence of such a link.

We have seen that there is variation in the acceptability judgments of sentences involving intervention effects; that is, there is variation in the perceived *strength* of intervention effects (cf. target responses in Figure 6.3). Furthermore, previous experimental work has shown that people differ in how prone they are to derive scalar implicatures (Noveck and Posada, 2003). The goal of Experiment 2 was to determine whether there is a relationship between the strength of intervention effects and the rate of derivation of scalar implicatures at an individual level. Such a relationship is expected under the SI theory, since not deriving the implicature should provide access to a grammatical parse of the sentence that would otherwise be considered ungrammatical due to the intervention effect.

6.3.1 Participants

54 participants (24 female) were recruited on Amazon Mechanical Turk and were paid $1.80 for their participation. Two participants were excluded from analysis for not being native speakers of English.

6.3.2 Procedure and materials

Experiment 2 involved two tasks: a Picture Selection Task (Roeper, 2007) was used to estimate the participants' rates of implicature derivation, and an Acceptability Judgment Task was used to estimate the strength of the intervention effects. The order of the two tasks was counterbalanced across participants.

6.3.2.1 Acceptability Judgment Task

This task was almost identical to Experiment 1 in terms of the instructions and materials. There were two important differences, mostly imposed by the need to make room for a second task. First, there was no [-NPI] condition: all sentences in this task contained the word *any*. This prevented the possibility of evaluating the *contribution* of the NPI through the log-ratios as above, but to compensate for that we did not ask participants to report overall judgments about the sentences, but rather to report judgments about the contribution of the word *any*. Second, only the polar environments from Experiment 1 were tested: the target intervention environment, and one UE and one DE environment.

The Acceptability Judgment Task consisted of 19 items, which were preceded by three example items. The first item had an unambiguously unlicensed occurrence of *any*; in the immediate feedback, participants were told that most people would judge this item low on a scale from one to five. The second example item had a licensed *any* in the restrictor of a universal quantifier; the feedback to participants following this item was that most people would judge this item high on a scale from one to five.[5] The third example item was harder to judge; it contained *any* in a non-monotonic environment, which we know elicits variable judgments across speakers. Participants were told that some people would judge this item low and others high on the scale, and that they should simply follow their intuitions. The three examples were then presented again as the first three items of the experiment. The remaining 16 items consisted of eight target items and eight control items, presented in a pseudo-randomized order for each participant.

Target items were similar to the corresponding items from Experiment 1. An example target item is repeated in (10).

(10) Monkey didn't give every rabbit any juice.

There were four control items with a licensed *any* in the scope of negation, as in (11a), hereafter referred to as 'good controls'. The other four control items contained an unlicensed NPI *any*, as in (11b), hereafter referred to as 'bad controls'.

(11) a. Rabbit didn't drink any tea.
b. Lion drank any coffee.

6.3.2.2 Picture Selection Task

In the Picture Selection Task of Experiment 2, participants saw two pictures on each trial with a very short introduction sentence at the top of the screen that provided a setting for the story, and

[5]This example item did not include the same licensor (the restrictor of 'every') as the test items (negation), to avoid the possibility that participants would simply repeat (an abstract form of) the judgment they were given at the outset.

read a sentence that they were told had been produced by a puppet named Raffie.[6] Participants were instructed to evaluate Raffie's sentence with respect to the pictures on the screen. They were told that Raffie's sentence would sometimes be applicable to only one of the two pictures (Picture 1 on the left or Picture 2 on the right), and sometimes to both of the pictures. For each test sentence, participants were asked to decide from three response options: Picture 1, Picture 2, Both Pictures. An example item can be seen in Figure 6.4.

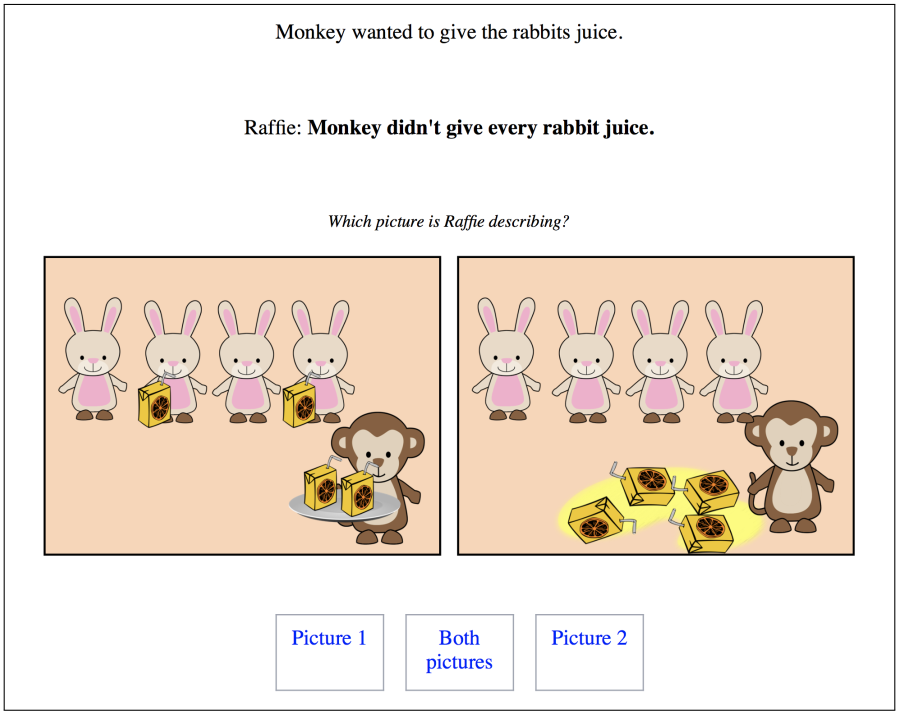


Figure 6.4: Experiment 2: An example of a Picture Selection Task trial

The Picture Selection Task consisted of 31 items. Participants were first presented with three examples. Raffie's sentence in the first example clearly matched Picture 1. Raffie's sentence in the second example clearly matched Picture 2. Raffie's sentence in the third example was harder to judge, and participants were told that some people would say it matched only one of the pictures, while others would say it matched both pictures, and they should simply follow their intuitions while doing the task. The three examples were then presented again as the first three items of the experiment. The remaining 28 items consisted of eight target items and 20 controls, presented in randomized order.

The target items in the Picture Selection Task were exactly parallel to those in the Acceptability Judgment Task, the only difference being that there was no NPI in the target sentences in the Picture Selection Task. An example target item from the Picture Selection Task is provided in Figure 6.4. In this example, the target sentence corresponds to (12a). Its logically stronger scalar alternative is in (12b). The scalar implicature of (12a) is thus the negation of (12b), i.e. the resulting interpretation of (12a) enriched with its scalar implicature is (12c).

[6]The materials for the study were made to be child-friendly, to allow for the possibility of a parallel child language acquisition study.

(12) a. Sentence: Monkey didn't give every rabbit juice.
b. Alternative: Monkey didn't give any rabbits juice.
c. Overall meaning with SI: Monkey gave some but not every rabbit juice.

In Picture 1, Monkey gave two of the four rabbits juice, and in Picture 2 Monkey dropped all of the juiceboxes and none of the rabbits got juice. Now, if the participant derived the implicature, they would understand the sentence as (12c) (by the negation of the logically stronger alternative *Monkey didn't give any rabbits juice*). They would thus opt for the response 'Picture 1'. We will refer to this type of response as a *some*-response (as some of the animals in the picture received juice). On the other hand, if the participant didn't derive the implicature, they would interpret the sentence literally, true both in the situation in which Monkey gave only some of the rabbits juice (Picture 1) and in the situation in which he gave none of them juice (Picture 2); in this case, the participant would opt for the response 'Both pictures'.[7] Based on their response on target items, we can thus evaluate whether the participant interpreted the sentence with or without the implicature.

In addition to the targets, participants also received 20 control sentences. Four had the same construction as the target items, but their meaning was compatible with only one of the pictures. An example of this type of control is provided in (13a), which was paired with a picture in which two out of four elephants got cheese, and a picture in which four out of four elephants got cheese. Participants also saw four positive ditransitive control sentences, as in (13b), four negative ditransitive sentences with the universal quantifier in direct object position, as in (13c), and four negative ditransitive sentences with a definite noun phrase in indirect object position and a mass noun in indirect object position, as in (13d). For half of the controls, the correct response was 'Picture 1', and for the other half the target response was 'Picture 2'. Finally, there were four sentences like (13e), which were compatible with both of the pictures they were presented with.

(13) a. Dog didn't give every elephant cheese.
b. Elephant gave every dog cake.
c. Giraffe didn't give Lion every strawberry.
d. Rabbit didn't give Cat tea.
e. Giraffe got every flower.

6.3.3 Results

6.3.3.1 Exclusions

Target sentences could be associated with two possible interpretations, one with the implicature and one without the implicature. There were therefore only two possible responses that participants could give: if the participant derived the implicature, they were expected to give a *some*-response (e.g., 'Picture 1' in Figure 6.4); if the participant didn't derive the implicature,

[7] As an anonymous reviewer points out, participants who selected the response 'Both pictures' might have been aware of both readings of the sentence: one with and one without the scalar implicature. For our purposes, however, it is not crucial to distinguish between participants who selected 'Both pictures' because they only accessed the implicature-less reading and those participants who selected 'Both pictures' because they had access to both the implicature- and implicature-less readings. What matters for us is that all of these participants in effect did have access to the implicature-less reading; under the implicature theory of intervention effects, all of these participants effectively had access to a grammatical parse of sentences containing an intervention configuration.

they were expected to select 'Both pictures'. Under no legitimate interpretation was a participant expected to select the image in which no animal gets the object in question (e.g., 'Picture 2' in Figure 6.4); we will hereafter refer to these illegitimate responses as *none*-responses. *None*-responses would only be possible if the universal quantifier *every NP* could scope above negation, but this reading is normally unattested. Only one participant opted exclusively for such answers on our target items, so we decided to exclude them from the analysis. The remaining participants gave very few *none*-responses (6 out of 392), and these responses were excluded from the analysis as well.

Participants also had to correctly answer at least 75% of the controls in the Picture Selection Task to be included in the data analysis. This led to the exclusion of two additional participants.

For the Acceptability Judgment Task, participants' individual responses had to be such that, schematically: mean(good controls) ≥ mean(target) ≥ mean(bad controls). This requirement led to the exclusion of nine more participants, which left us with a total of 40 participants for analysis.

The 40 participants retained for the analysis responded with an average of 94.3% accuracy on control items in the Picture Selection Task. As for the controls in the Acceptability Judgment Task, the mean judgment for the good controls was 4.94 ($SD = 0.18$), and the mean judgment for the bad controls was 1.44 ($SD = 0.6$).[8]

6.3.3.2 SI derivation

On the Picture Selection Task targets, the proportion of responses consistent with an implicature (SI+ responses) was 0.56, with a standard deviation of 0.48.

6.3.3.3 Intervention measure

In the Acceptability Judgment Task, the critical target sentences received an average rating of 3.76 ($SD = 1.02$), which differed significantly from ratings for the good controls ($t(39) = -7.33$, $p < .001$) and the bad controls ($t(39) = 12.8$, $p < .001$).

6.3.3.4 Correlation

The main goal of Experiment 2 was to assess whether there was a correlation between implicature derivation and intervention effects.[9] First, we normalized each participant's responses on target items in the Acceptability Judgment Task by scaling each target response within the particular participant's extreme judgments of the good and bad controls, as in: $\frac{\text{target-mean}_p(\text{bad})}{\text{mean}_p(\text{good})\text{-mean}_p(\text{bad})}$, where mean_p represents the mean of responses in a particular condition for a given participant p. This normalization corrects for different uses of the response scale by matching extreme values across participants. In Figure 6.5 below, each participant is represented at the height corresponding to the mean of these normalized responses for that given participant $\frac{\text{mean}_p(\text{target})\text{-mean}_p(\text{bad})}{\text{mean}_p(\text{good})\text{-mean}_p(\text{bad})}$. This is a value between 0 (intervention effects are as bad as bad controls) and 1 (intervention effects are as good as good controls), since we excluded participants who

[8] Means and standard deviations (here and elsewhere) are calculated on within-participant means.

[9] We present all the results collapsed independently of which block occurred first: comparisons of mixed effects linear regression models (including random by-participant intercepts) with and without order of the two tasks as a fixed effect showed no effect of task order on intervention effects ($\chi^2(1) < 0.001$, $p = .98$). Similarly, the comparison of mixed effects logit models with and without the order of the two tasks as a fixed effect showed no effect of the order of the task on implicature derivation ($\chi^2(1) = 0.004$, $p = .94$).

did not satisfy $\text{mean}_p(\text{good}) \geq \text{mean}_p(\text{target}) \geq \text{mean}_p(\text{bad})$. To assess the correlation between the strength of intervention effects so represented and implicature derivation, we computed an implicature index for each participant: the proportion of SI+ responses to targets in the Picture Selection Task by that given participant. This implicature index is represented on the x-axis of Figure 6.5.

The first thing to note about this graph is that there is a wide range of perceived strength of intervention effects (participants are distributed all along the range of the y-axis), which replicates the finding from Experiment 1 that there is a lot of variation in the perception of intervention effects. The implicature index also shows variability, but in the form of bimodality (participants fall either to the far left or to the far right of the graph).

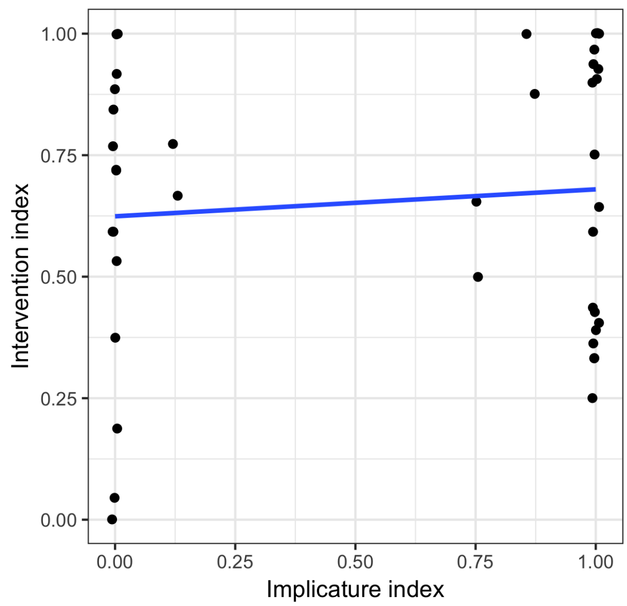


Figure 6.5: Results from Experiment 2: Individuals as a function of responses to implicature and intervention targets

Turning to the correlation, mixed effects linear regression models were fitted to the normalized responses to the Acceptability Judgment Task targets, with mean response to the implicature targets as a fixed effect, and random by-participant intercepts. Comparisons with the models containing only the random by-participant intercepts revealed no significant effect of implicature response on the intervention judgments ($\chi^2(1) = 0.35, p = .55$). Experiment 2 thus reveals no correlation between an individual tendency to derive implicatures and individual perception of intervention effects.[10]

[10] One could investigate other predictions, more specific to the scalar implicature account. The standard view holds that the derivation of scalar implicatures requires extra processing time (see Bott and Noveck, 2004 and much subsequent work, see Cremers and Chemla, 2014 for similar results on the relevant *indirect* scalar implicatures, but also Romoli and Schwarz, 2015 for an opposing view). If this is correct, then one would expect that low judgments of intervention effect sentences would require extra time in the Acceptability Judgment Task. Although this is borne out ($\chi^2(1) = 3.67, p = .05$), such an effect is not stronger than a similar effect under which lower judgments of ungrammatical sentences in general take more time; the interaction between the two types of effects is not statistically significant ($\chi^2(1) = 0.99, p = .32$).

6.3.4 Discussion

In this experiment, we measured participants' propensity to derive SIs and their sensitivity to intervention effects. While we observed variability in both, there was no observed correlation between the two indices.

6.4 Experiment 3

Experiment 3 had the same goal and general design as Experiment 2, except that it used a different task to estimate participants' propensity to derive scalar implicatures. In brief, we will see that Experiment 3 replicates the results of Experiment 2, revealing an absence of correlation between intervention effects and implicature derivation.

Participants in Experiment 2 were very consistent in the Picture Selection Task with respect to whether their responses were consistent with the implicature or consistent with an interpretation without the implicature. Given that intervention effects do not show such a robust bimodality, it is not surprising that SIs and intervention effects are not well correlated. However, the bimodality of the SI results may have been an artifact of the particular task we used to test for implicatures; thus in Experiment 3 we opted for a similar design but an alternative method for testing for SI derivation: the Covered Picture Task (see Huang et al. 2013).

6.4.1 Participants

55 participants (27 female) were recruited through Amazon Mechanical Turk and were paid $1.80 for their participation. All participants reported English as their native language.

6.4.2 Procedure and materials

The procedure was the same as in Experiment 2, comprising two tasks. In addition to the Acceptability Judgment Task, participants completed a Covered Picture Task. In this task, they saw two pictures on the screen, one visible and the other covered, and a short introduction sentence at the top of the screen to introduce the story, followed by a target sentence produced by a puppet called Raffie. The participants were instructed to evaluate the visible picture with respect to Raffie's sentence. They were told that Raffie's sentence could apply to only one of the pictures — either the visible or the covered one. The participants had to click on the visible picture if they thought that Raffie's sentence was describing this picture, and on the covered picture if they thought that Raffie's sentence was describing a different picture from the visible one. An example of a target item is provided in Figure 6.6.

The visible picture that appeared with the target items was one that was incompatible with the derivation of the implicature. In the visible picture in Figure 6.6, we can see that none of the rabbits got juice. The participant would thus opt for the visible picture if they had access to the implicature-less reading, and for the covered picture if they expected, based on the target sentence, that at least some of the rabbits got juice, which is the reading of the target sentence enriched with an implicature, as in (12c). As in the Picture Selection Task of Experiment 2, we can evaluate based on responses to the target items whether the participant has interpreted the sentence with or without the implicature.

The Covered Picture Task contained 29 items. Participants were first presented with three examples. The first one involved a correct description of the visible picture; the immediate feedback to participants stated that in that case most participants would select the visible picture. The second example involved an incorrect description of the visible picture; here the feedback

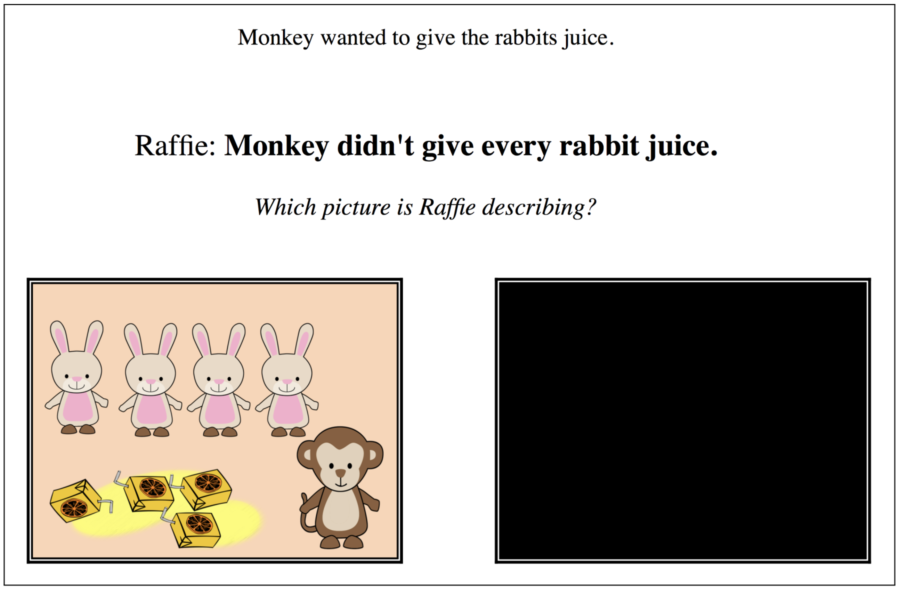


Figure 6.6: Experiment 3: An example of a Covered Picture Task trial

to participants was that most participants would select the covered picture. The third example was harder to judge; for this item, the participants were told that some people would select the visible picture and others the covered picture, and that they should follow their intuitions while doing the task. These three examples were then presented again as the first three items of the experiment. The remaining 26 items consisted of eight target items and 18 controls, presented in randomized order.

Among the 18 control items, 10 were compatible with the visible picture, and eight were not. They were very similar to the controls in the Picture Selection Task of Experiment 2 (but adapted for the Covered Picture Task).

6.4.3 Results

6.4.3.1 Exclusions

The exclusion criteria were the same as in Experiment 2, leading to the exclusion of 9 participants from analysis. The 46 participants retained for the analysis responded with an average of 92.1% accuracy on control items in the Covered Picture Task. In the Acceptability Judgment Task, the mean judgment for good controls was 4.84 (SD = 0.42), and the mean judgment for bad controls was 1.62 (SD = 0.72).

6.4.3.2 SI derivation

In the Covered Picture Task, the proportion of SI+ responses to the targets was 0.22, with a standard deviation of 0.31.[11]

[11] An anonymous reviewer points out that the estimate of implicature derivation in a covered picture paradigm may be noisy, as it incorporates biases across individuals, such as a bias against selecting the covered picture. To the extent that such a bias is not correlated with an individual participant's implicature derivation, with sufficient power it should not influence the main effect we are interested in.

6.4.3.3 Intervention effects

In the Acceptability Judgment Task, the critical targets received an average rating of 3.73 (SD = 1.04), which differed significantly from good controls ($t(45) = -8.39$, $p < .001$) and bad controls ($t(45) = 13.76$, $p < .001$).

6.4.3.4 Correlation

As in Experiment 2, we calculated each participant's intervention index, represented on the y-axis of Figure 6.7, and the proportion of implicature derivation, represented on the x-axis.[12] The closer the value on the x-axis to 1, the more implicatures the participant derived in the course of the Covered Picture Task; the closer the value on the y-axis to 1, the greater the judgments provided by the participant in response to the intervention targets in the Acceptability Judgment Task.

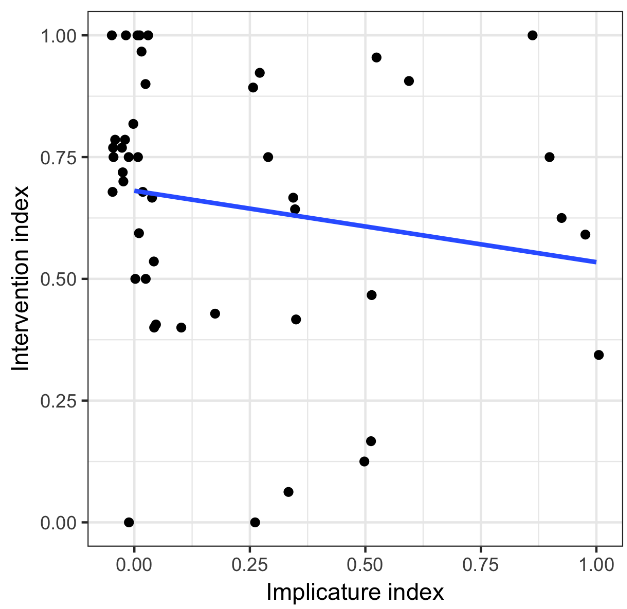


Figure 6.7: Results from Experiment 3: Individuals as a function of responses to implicature and intervention targets

A mixed effects linear regression model was fitted to the normalized[13] responses to the Acceptability Judgment Task targets, with mean response to the implicature targets as a fixed effect and random by-participant intercepts. A comparison with the model containing only the random by-participant intercepts revealed no significant effect of covered picture implicature responses on the intervention judgments ($\chi^2(1) = 1.18$, $p = .28$).

6.4.4 Discussion

The results of Experiment 3 replicate those of Experiment 2: whether one uses a Picture Selection Task or a Covered Picture Task to estimate participants' rates of implicature derivation, we

[12]Similar statistical methods as in Experiment 2 showed no effect of task order on either judgments of intervention effects ($\chi^2(1) = 0.86$, $p = .35$) or implicature derivation ($\chi^2(1) = 0.46$, $p = .5$). We therefore present all the results collapsed.

[13]Target responses were normalized in the same way as in Experiment 2.

find that this estimate does not predict participants' acceptability judgments of intervention sentences.

6.5 Experiment 4: Training study

Experiments 2 and 3 revealed no correlation between participants' spontaneous rate of implicature derivation and their spontaneous grammaticality judgments of intervention sentences. In Experiment 4, we drew on previous experimental work revealing that people can be trained to derive implicatures or not to derive implicatures (see, for example, Noveck and Posada 2003; Bott and Noveck 2004; Chemla et al. 2017). We investigated whether training participants either to derive or not to derive implicatures would influence the strength of intervention effects, comparing the performance of a group that received training consistent with implicature derivation, with that of a group that received training inconsistent with implicature derivation.

6.5.1 Participants

54 participants (25 female) were recruited through Amazon Mechanical Turk and were paid $1.80 for their participation. All participants reported English as their native language.

6.5.2 Procedure and materials

The experiment was divided into two parts. The first task was designed to train half of the participants to derive SIs and the other half to not derive SIs. In the second part, participants received an Acceptability Judgment Task (designed to measure intervention effects), as in Experiments 2 and 3.

In the implicature training task, participants completed a Truth Value Judgment Task (TVJT) (Crain and Thornton, 1998) (see Figure 6.8) in which they received immediate feedback after each trial. Participants saw a series of pictures with a short sentence at the top of the screen that introduced the story, followed by a target sentence produced by Raffie. Participants were instructed to evaluate the picture with respect to Raffie's sentence: they had to decide whether the descriptions were right or wrong. The response options were binary: the participants could either choose a 'yes' or a 'no' response.

If the participant derived the implicature, they would consider Raffie's sentence false in a situation in which none of the lions got ice cream (Figure 6.8), and would therefore select the 'no' response. On the other hand, if the participant didn't derive the implicature, they would consider Raffie's sentence to be an acceptable description of the picture, and would select the 'yes' response. What distinguished the group that was trained to derive implicatures (SI+ training group) from the group that was trained not to derive implicatures (SI- training group) was the feedback after the participant had provided a response on each of the target items. For example, for the experimental item in Figure 6.8, participants in the SI- training group received the feedback in (14), and participants in the SI+ training group received the feedback in (15).

(14) Feedback for SI- training group

a. 'Yes' response: *Correct!*

b. 'No' response: *Raffie was actually right — she said Bear didn't give every lion ice cream, and he didn't!*

(15) Feedback for SI+ training group

Figure 6.8: Experiment 4: An example of an implicature training TVJT trial

a. ‘Yes’ response: *Raffie was actually wrong — she said Bear didn’t give every lion ice cream, but none of the lions got ice cream!*
b. ‘No’ response: *Correct!*

The task consisted of 28 items. Participants first saw two example items, the first a clearly true target and the second a clearly false target. For both example items feedback was provided about which response was correct. The example items were then presented again as the first two items of the implicature training TVJT. The remaining 26 items consisted of eight target items and 18 control items, presented in randomized order.

The target sentences on the implicature training TVJT were the same as those from Experiments 2 and 3. The pictures that accompanied these sentences were incompatible with the derivation of implicatures. The control items in this experiment were similar to those in Experiments 2 and 3. Of the 18 control items, 10 items had a clear ‘yes’ target, and eight items had a clear ‘no’ target. The main difference between these controls and those of the previous experiments was that the control trials in this experiment contained feedback, so as to be parallel with the target trials.

6.5.3 Results

6.5.3.1 Exclusions

The exclusion criteria were the same as in Experiments 2 and 3, leading to the exclusion of eight participants from analysis. The remaining 46 participants responded with an average of 95% accuracy on control items in the implicature training TVJT. In the Acceptability Judgment Task, the average judgment for good controls was 4.87 ($SD = 0.34$), and the average judgment for bad controls was 1.85 ($SD = 0.99$).

6.5.3.2 Training results

In the implicature training TVJT, the proportion of SI+ responses from the group of participants that received the SI+ training was 0.58 ($SD = 0.25$), while the group that received the SI- training gave no SI+ responses. This suggests that (i) the implicature-less response is salient from the start but (ii) the training is effective. To see this in more detail, Figure 6.9 displays the proportion of SI+ responses to the critical cases, arranged by the order in which they were presented. This graph allows us to visualize the effect of the SI+ training: all of the participants started off without implicatures, but they became more and more likely to derive the implicature with each subsequent SI+ training target.[14]

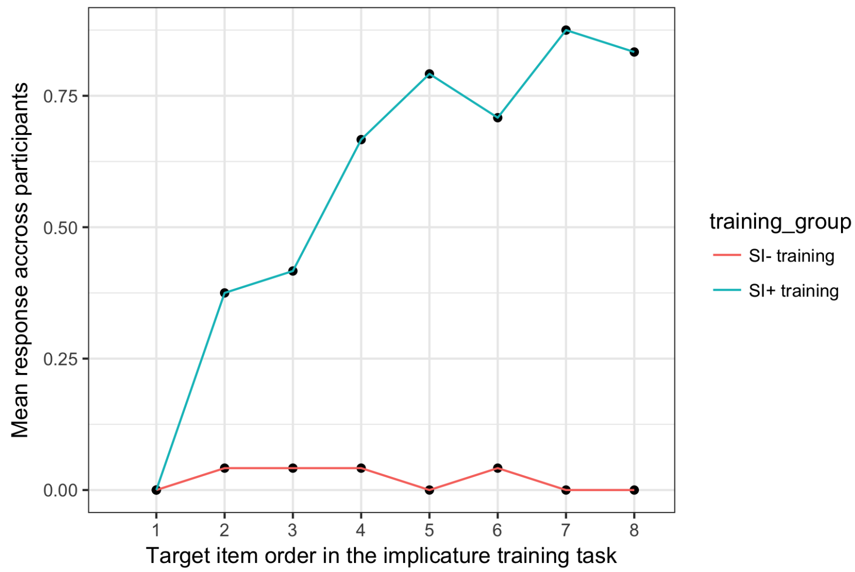


Figure 6.9: Experiment 4: Proportion of SI+ responses to implicature targets for each training group, by order of target appearance

6.5.3.3 Intervention effects

In the Acceptability Judgment Task, participants who received the SI+ training gave the targets an average rating of 3.91 ($SD = 0.84$), which differed significantly from good controls ($M = 4.93$, $SD = 0.21$, $t(23) = -6.33$, $p < .001$), and from bad controls ($M = 1.79$, $SD = 0.9$, $t(23) = 9.42$, $p < .001$). Participants who received the SI- training gave an average rating of 3.71 ($SD = 0.98$), which differed significantly from good controls ($M = 4.81$, $SD = 0.4$, $t(21) = -5.78$, $p < .001$), and from bad controls ($M = 1.92$, $SD = 1.13$, $t(21) = 6.91$, $p < .001$).

A mixed effects linear regression model was fitted to the normalized[15] responses to the Acceptability Judgment Task targets, with the training received as a fixed effect, and random by-participant intercepts. A comparison with the model containing only the random by-participant intercepts revealed no significant effect of training received in the implicature training TVJT on subsequent acceptability judgments on the intervention items in the Acceptability Judgment Task ($\chi^2(1) = 0.78$, $p = .38$).

[14] An anonymous reviewer asks whether a similar increase in implicatures was observed in Experiments 2 and 3, which did not include a training component. We observed no such increase in implicatures across subsequent trials in Experiments 2 and 3, suggesting further that the SI+ training administered in Experiment 4 was indeed effective at encouraging participants to derive implicatures.

[15] Target responses were normalized in the same way as in Experiments 2 and 3.

6.5.4 Discussion

The training administered in Experiment 4 succeeded in leading participants to either derive implicatures or to not derive implicatures. However, this training effect did not subsequently affect their acceptability judgments of the intervention targets. While the SI theory might have led us to expect that the SI+ group would judge the intervention targets worse than the SI- group, this prediction was not borne out.

6.6 Experiment 5: Repair strategy study

Experiments 2-4 revealed no relationship between the rate of implicature derivation and the strength of intervention effects. This is surprising under the hypothesis that intervention effects are caused by the presence of scalar implicatures. The failure to observe this correlation could come from the fact that we measured SI derivation and strength of intervention at different points in time, while participants were carrying out different tasks. Hence, the SI derivation rates we measured may not be telling us much about whether participants derived implicatures *when they were providing judgments about the intervention effects.*

In Experiment 5, we thus attempted to measure the joint effects of implicature derivation and acceptability at a single point in time. The idea was the following. It is possible to assign a meaning to some sentences despite the fact that they are ungrammatical: we often and easily apply systematic 'repair strategies' to non-grammatical sentences. The goal of Experiment 5 was to detect whether the suppression of the implicature is an available repair strategy for sentences containing intervention configurations. If the source of the ungrammaticality of the NPI in intervention configurations is the scalar implicature, a natural repair strategy would be to ignore the implicature and thereby allow the NPI to be licensed.

6.6.1 Participants

52 participants (24 female) were recruited through Amazon Mechanical Turk and were paid $1.80 for their participation. All participants reported English as their native language.

6.6.2 Procedure and materials

The experiment consisted of a single task, which was equivalent to the Picture Selection Task from Experiment 2. Participants were randomly assigned to one of two groups, which differed only in the target sentences they saw. The 'regular' group saw the standard 8 [-NPI] target sentences from Experiment 2 (Appendix B.2, (6)). These sentences did not contain an intervention configuration, and therefore there was no need for any repair strategies. The 'intervention' group saw almost the same 8 target sentences, except that these contained the NPI *any* in an intervention configuration. If participants in the intervention group detected an intervention effect, they were expected to *repair* the sentence in order to carry out the task. For example, the intervention group might see a target like that in Figure 6.10, while the regular group would see a target like Figure 6.4 from Experiment 2 (which is identical except that it does not include the NPI *any*).

In addition to the critical targets, both groups received the same 20 control items from the Picture Selection Task of Experiment 2 (Appendix B.2, (7)).

The intervention group (but not the regular group) was also asked to fill out a small questionnaire at the end of the experiment. In this questionnaire, the participants were asked to motivate their response choice on the target items: this questionnaire was there to make sure

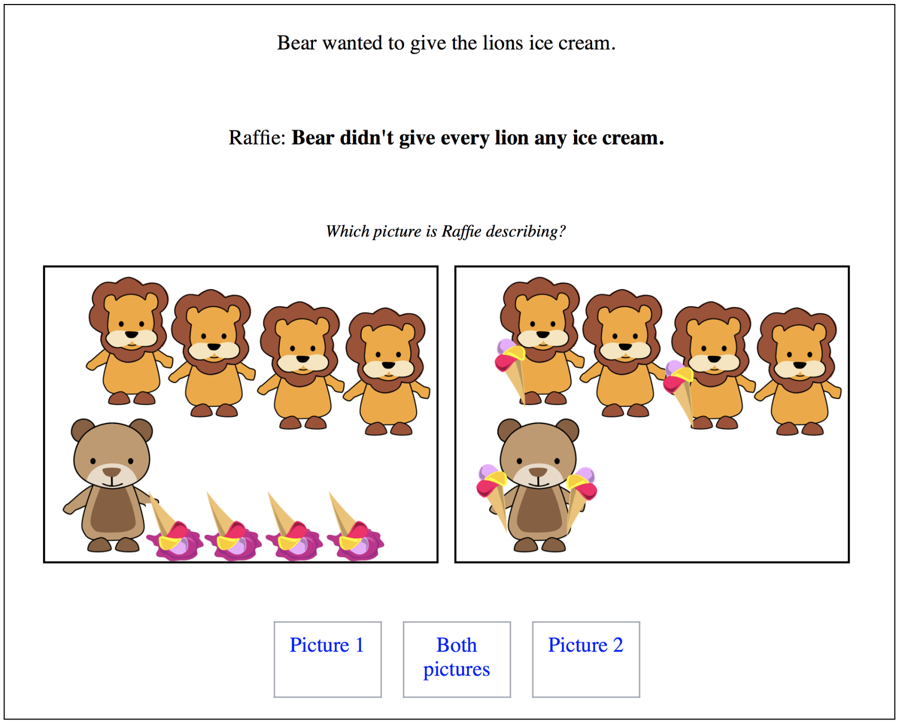


Figure 6.10: Experiment 5: An example of a target item administered to the intervention group

that participants had actually assigned a meaning to the target sentences, and had not simply responded at random.

6.6.3 Repair strategies

Before discussing the results, let us consider some possible repair strategies for a sentence like (16).

(16) *Bear didn't give every lion any ice cream.

One possibility would be to simply drop the NPI, and interpret the sentence as in (17a). Another possibility would be to reinterpret the universal quantifier as a definite description, as in (17b). Yet a third possibility would be to assign wide scope to the universally quantified noun phrase, above negation, as in (17c). The question of interest is whether there is a fourth possibility: to interpret the sentence as it is, *without the implicature*, as in (17d).

(17) a. Bear didn't give every lion ice cream.
b. Bear didn't give the lions any ice cream.
c. Every lion is such that Bear didn't give him ice cream.
d. Bear didn't give every (and possibly[16] he didn't give any) lion ice cream.

What would be the expected responses on the Picture Selection Task under each of the

[16]This paraphrase is in fact too strong; the 'possibly' should not make one think that the intended reading is one that is false when the possibility is out of the question. An appropriate paraphrase would be 'Bear didn't give every lion ice cream or he didn't give any lion ice cream', but again, without the ignorance inference now attached to the disjunction (instead of to 'possibly').

above repair strategies? Table 6.2 summarizes possible repair strategies for (16) and their predictions, which we will now spell out. We will continue to refer to selections of pictures in which no animals got the object in question (e.g., a Picture 1 response in Figure 6.10) as *none*-responses,[17] and to selections of pictures in which some animals got the object in question as *some*-responses (e.g., a Picture 2 response in Figure 6.10).

Participants who adopt the strategy that leads to the interpretation in (17a), i.e. dropping the NPI, might respond in one of two ways, depending on whether they compute an indirect scalar implicature from (17a). If they compute the implicature *Bear gave some of the lions ice cream*, they are expected to give a *some*-response. If they do not compute an implicature from (17a), they are expected to select 'Both pictures', since the implicature-less reading of (17a) is compatible with both pictures. Notice that if dropping the NPI is the strategy adopted by the intervention group, we may expect to see parallel performance in the two groups, since the reanalyzed (17a) is identical to the target sentences that the regular group is presented with.

Participants who employ the strategy of reinterpreting the universal quantifier as a definite description, as in (17b), or of giving the universal wide scope, as in (17c), would be expected to give a *none*-response. If the intervention group makes use of either of these two strategies, we expect to see a greater proportion of *none*-responses in the intervention group than in the regular group.

Finally, participants who employ the strategy of implicature suppression as in (17d) are expected to select 'Both pictures'. Since none of the above repair strategies lead to an increase in 'Both pictures' responses, an increase of such responses in the intervention group could be interpreted as the propensity to employ that strategy (to be discussed further below).

Table 6.2: Summary of repair strategies and the predicted responses

Repair strategy	Post-repair interpretation	Prediction
Drop the NPI	(17a)	Same result pattern as in Experiment 2
Replace the universal quantifier with a definite description	(17b)	Increase of *none*-responses
Assign wide scope to the universal quantifier	(17c)	Increase of *none*-responses
Do not derive an implicature	(17d)	Increase of 'Both pictures' responses

6.6.4 Results

6.6.4.1 Exclusions

As in the earlier experiments, participants had to correctly answer at least 75% of the controls in order to be included in the data analysis. One participant failed to do so and was excluded from the analysis. The mean accuracy on controls for the remaining 51 participants was 94%.

[17] Note that these responses should now be perfectly legitimate for the 'intervention' group in the experiment.

6.6.4.2 Targets

Figure 6.11 displays the counts of the different response types for the implicature targets from the intervention group and the regular group. Recall that a 'Both pictures' response corresponds to a non-implicature interpretation of the target sentence. To determine whether group (intervention vs. regular) had an effect on the type of response to the implicature targets, we recoded the responses in binary terms (as 'Both pictures' and 'Other'). We then fitted a mixed effects logistic regression model on these responses, with group as a fixed effect and random by-participant intercepts. A comparison of this model with a reduced model without the group fixed effect revealed a significant effect of group on response type ($\chi^2(1) = 6.05, p = .01$), with more 'Both pictures' (non-implicature) responses in the intervention group than in the regular group.

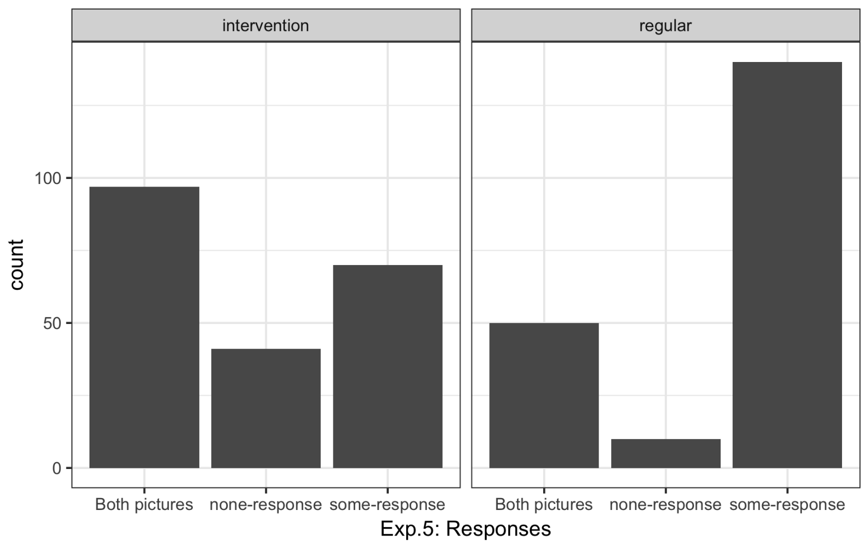


Figure 6.11: Experiment 5: Distribution of responses to implicature targets from the intervention and the regular group

6.6.5 Discussion

One interpretation of these results is that they are consistent with the SI theory: if scalar implicatures cause intervention effects, it is natural to expect that when faced with an NPI in an intervention configuration, participants may choose to repair the sentence by cancelling the derivation of the scalar implicature.

However, there are two alternative explanations for this effect. First, note that there is evidence that at least some of our participants repaired the target sentences by replacing the universal quantifier with a definite description or by means of wide-scoping the universal quantifier: this is suggested by the increase in the amount of *none*-responses in the intervention group compared to the regular group. As these strategies seem to be available to some of our participants, the participants who selected 'Both pictures' could have entertained not one but *two* possible repair strategies simultaneously, namely, one of the strategies leading to a *none*-response, in combination with the strategy of dropping the NPI. Such participants might simply not have known which of the two interpretations should be chosen. For example, such a participant might think that it is as probable that Raffie wanted to use an indefinite instead of the NPI, and derive the implicature (which would make the sentence compatible with Picture 2 of Figure 6.10), as it is that the universal quantifier should be assigned wide scope with respect to negation (which would make Raffie's sentence compatible with Picture 1 of Figure 6.10). The participants could therefore be unsure which of the two meanings Raffie intended to convey, which is why they might opt for 'Both pictures' more often in the intervention group than in the regular group.

An argument against this explanation comes from the questionnaire that the intervention group was asked to complete at the end of the experiment. The explanations that participants gave for their picture selections can be found in Appendix B.5.2. These explanations confirm that for the most part, participants were not merely responding at random to the target items, but rather repaired the sentences to assign some meaning to them. The crucial question for us was whether participants who selected 'Both pictures' might report confusion or uncertainty about the intended meaning of the sentence, and whether this uncertainty could have driven their picture selection. Upon close inspection, however, few responses were indicative of such confusion, with only one of the 10 'Both pictures' justifications referring to an ambiguity (i.e. *The sentence is a bit ambiguous and could either mean that the elephant didn't give any tea to the giraffes, or that he didn't give tea to all the giraffes*). The questionnaire responses therefore speak against an alternative explanation of the data according to which people were unsure about the intended meaning of the sentence due to the existence of more than one repair strategy.

While we do not find the multiple repair strategy explanation very strong in light of the questionnaire data, it is worth noting that we cannot at present completely rule it out either. Even when participants' explanations for 'Both pictures' selections are compatible with the non-implicature reading (e.g., *In both pictures not every giraffe has tea, so both are correct*), we cannot definitively know whether they chose 'Both pictures' because they did not derive the implicature, or because there were multiple repair strategies that would render the sentence true as soon as not all of the giraffes got tea.

A second alternative explanation is that participants in the repair group were under more cognitive load, which is why they derived fewer implicatures. Previous experimental work has shown that in a dual task setting people tend to prefer readings without scalar implicatures (De Neys and Schaeken, 2007; Marty and Chemla, 2013). Facing a broken sentence (with an intervention effect) and perhaps trying to interpret it could require additional cognitive load leading to fewer implicatures, not because implicature suppression is a possible repair strategy, but because participants are expending effort to repair the sentence.

Overall, one might have considered this experiment to hold the greatest chance of yielding an observable correlation between intervention effects and SIs, because it allowed both to be measured at a single point in time. Compared to the other experiments, this one yields the results most compatible with a positive conclusion in favor of a correlation, although there remain alternative explanations that require further investigation.

6.7 General discussion

The current studies provide the first quantitative measurement of the strength of intervention effects. Throughout all of the experiments, we observe that intervention effects are real, but not as strong as one might have expected.

A more specific goal of the present experiments was to use such measurements to evaluate the relationship between scalar implicature derivation and the presence of intervention effects in NPI licensing. The SI theory posits that intervention effects are caused by scalar implicatures. This raises a tension between the optionality of SIs and the categorical judgments reported in the theoretical literature on intervention effects. Experiment 1 suggests that one might be able to resolve this tension. In this experiment (as well as in subsequent experiments), intervention sentences were found to be judged better than different kinds of sentences with unlicensed NPIs, and there was great variability in how ungrammatical (if ungrammatical at all) people considered sentences with intervention configurations. This response pattern to intervention sentences goes well with the hypothesis that intervention effects are caused by scalar implicatures, as it could be that it is precisely the optionality of implicature derivation that leads to the variability in judgments of intervention sentences.

In Experiments 2 and 3, we thus investigated a more precise consequence of the implicature theory, whereby intervention effects should be contingent on the derivation of implicatures. But we observed no correlation between an individual's rate of implicature derivation and their sensitivity to intervention effects. In Experiment 4 we found that training people to derive or not to derive implicatures did not have an influence on their subsequent judgments of intervention effects. The results of Experiment 5, on the other hand, do suggest that when forced to assign a meaning to a sentence with an intervention configuration, participants behaved as though they had made it possible by blocking the derivation of scalar implicatures.

The present experiments thus paint a mixed empirical landscape. While we do not observe a relationship between judgments of intervention sentences and implicature derivation in Experiments 2-4, we do observe such a relationship in Experiment 5. Hence these results must be interpreted with care; if the implicature theory is on the right track, the absence of an effect in Experiments 2-4 must be explained. Note also that there is an important methodological difference between the experiments in which we observe null effects and the experiment in which we observe the expected effect, which may be quite relevant to explaining the findings.

In particular, Experiments 2-4 show that, if indeed intervention effects are caused by scalar implicatures, this will not manifest itself in a simple correlation between a task that tries to provide an individual's scalar implicature index, and a task that tries to provide an index of the individual's sensitivity to intervention effects. There could be multiple reasons for this; one possibility is that, as the two tasks are very different and implemented at different points in time, scalar implicature derivation in one task does not predict scalar implicature derivation later, in the other task.

Experiment 5, on the other hand, provided a *simultaneous* estimate of intervention effects and implicature derivation. This experiment is thus methodologically more powerful in terms of capturing the link between implicatures and intervention effects, if there indeed is one. These results must be interpreted with care, especially in light of the preceding absence of similar results in the previous experiments, but they are the best evidence of a link between implicatures and intervention effects.

Before closing, it is worth repeating that our study focused on testing a specific family of theories of intervention effects in NPI licensing, but there exist alternative accounts that do not appeal to scalar implicatures. For instance, Beck (2006) and Guerzoni (2006) relate the phenomenon to intervention effects in the domain of wh-words, where certain quantificational elements and operators like disjunction and conjunction can be seen to disrupt the licensing of wh-words as in (18) (examples adapted from Guerzoni, 2006).

(18) a. *Which book did which student and Mary read?
b. *Which book did which student or Mary read?

While reviewing these alternative theories would take us beyond the scope of the present paper, we would point out that the experiments reported here still offer a very relevant challenge for alternative accounts, namely how to explain the variability in judgments of intervention sentences and the fact that they are judged to be better than sentences containing unlicensed NPIs. Building on the present study, one prediction of accounts like Beck (2006) and Guerzoni (2006) that could be investigated in future research is that we might observe a similar variability in the grammaticality judgments of intervention effects in wh-licensing. From where we stand, it's unclear how these alternative accounts would explain the variability observed in our study, as well as any variability one might observe in judgments of intervention effects in wh-licensing. An ideal theory of intervention effects should seek to explain such variability, and future work could seek evidence for the connection with wh-intervention in the same way as we have done here for the SI theory.

Of course, it could always be that intervention effects and the observed variability are due to non-grammatical factors. For instance, the intervention configurations may impose a particular burden on participants, making grammatical sentences resemble ungrammatical ones to some participants (see Gibson, 1991, 1998 for related ideas about center embedding configurations or Sprouse et al., 2012 for related investigations on island effects).

Intervention could also be a real grammatical phenomenon, but one whose *variability* is due to individuals' limitations in independent processing abilities, e.g., working memory constraints. Even if such a scenario would take much of the task of explaining intervention out of the linguist's hands, it would be necessary to spell out and fully test such a proposal, in order to obtain a proper and complete understanding of intervention effects.

6.8 Appendix

6.8.1 Consent form and instructions

Consent form

Thank you for your participation.

We will be asking for your opinions about some English sentences. The material was designed for preschool-aged children; here, we are collecting adults' judgments for comparison. The task will take you about 15 minutes in total to complete.

Risks, benefits, confidentiality

There are no known risks or benefits associated with your participation in the study. All responses will be saved anonymously. When you are ready to begin, please click on the link below.

→ Click here to continue to the first task.

Figure 6.12: Consent form preceding each experiment

Instructions

In this task, you will see English sentences produced by Zap, an alien learning English, that are about animal characters: Rabbit, Lion, Monkey, Bear, Dog, Cat, Giraffe, and Elephant. Your task is to decide how Zap's sentences sound on a scale from 1 to 5, with 1 being completely odd and 5 being completely okay. Check the button and click on the link below to see some examples.

☐ I have read and understood the instructions.

→ Click here to see an example.

Figure 6.13: Experiment 1: Instructions

Instructions

In this task, you will see a series of short stories, each accompanied by two cartoon images. After each story, a puppet named Raffie will produce a sentence about what happened at the end of the story. Your task is to decide which of the two pictures matches Raffie's description. Sometimes Raffie's sentence will apply to only one of the pictures, and sometimes it will apply to both pictures. Note that Picture 1 will always refer to the picture on the left, and Picture 2 will always refer to the picture on the right. Check the button and click on the link below to see some examples.

☐ I have read and understood the instructions.

→ Click here to see an example.

Figure 6.14: Experiments 2 and 5: Picture Selection Task instructions

First part - Instructions

There will be two parts. In this first part, you will see a series of short stories, each accompanied by two images. One of them will be visible; the other is hidden by a black box so you cannot see it. At the end of each story, a puppet named Raffie will produce a sentence about what happened in the story. Your task is to decide which picture matches Raffie's description. If you think Raffie's sentence can apply to the visible image, click on it. Otherwise, assume that a better picture is hidden by the black box, and click on that one. Check the button and click on the link below to see some examples.

☐ I have read and understood the instructions.

→ Click here to see an example.

Figure 6.15: Experiment 3: Covered Picture Task instructions

First part - Instructions

There will be two parts. In this first part, you will see a series of short stories illustrated with cartoon images. At the end of each story, a puppet named Raffie will produce a sentence about what happened in the story. Your task is to decide whether Raffie is right or not. Check the button and click on the link below to see some examples.

☐ I have read and understood the instructions.

→ Click here to see an example.

Figure 6.16: Experiment 4: Implicature training TVJT instructions

Second part - Instructions

You have now completed the first part.

In the second part, you will see English sentences containing the word 'any', produced by Zap, an alien learning English. Your task is to decide how the word 'any' sounds in Zap's sentences, on a scale from 1 to 5, with 1 being completely odd and 5 being completely okay. Check the button and click on the link below to see some examples.

☐ I have read and understood the instructions.

→ Click here to see an example.

Figure 6.17: Experiments 2, 3, and 4: Acceptability Judgment Task instructions

6.8.2 Stimuli

Experiment 1

(19) **Example items:** Monkey poured the elephants some juice. Monkey is eating soup yet.

(20) [+**NPI**] **non-target items:** Cat didn't give the giraffes any coffee. ; Dog didn't give the bears any soup. ; Elephant didn't give the lions any honey. ; Rabbit didn't give the monkeys any water. ; Every giraffe who didn't have coffee drank any tea. ; Every rabbit who didn't have tea drank any water. ; Every bear who didn't have ice cream ate any chocolate. ; Every cat who didn't have chocolate ate any pie. ; Exactly two giraffes had any ice cream. ; Exactly two cats had any cheese. ; Exactly two lions drank any coffee. ; Exactly two elephants ate any cake. ; Giraffe had any ice cream. ; Cat had any cheese. ; Lion drank any coffee. ; Elephant ate any cake. ; Every giraffe who was hungry ate any chocolate. ; Every rabbit who was hungry ate any pie. ; Every bear who was thirsty drank any water. ; Every cat who was thirsty drank any juice.

(21) [+**NPI**] **target items:** Rabbit didn't give every monkey any soup. ; Dog didn't give every cat any water. ; Giraffe didn't give every elephant any coffee. ; Lion didn't give every bear any honey. ; Bear didn't give every giraffe any ice cream. ; Cat didn't give every rabbit any tea. ; Dog didn't give every lion any pie. ; Monkey didn't give every elephant any juice. ; Monkey didn't give every rabbit any juice. ; Cat didn't give every dog any pie. ; Elephant didn't give every giraffe any tea. ; Bear didn't give every lion any ice cream. ; Giraffe didn't give every bear any honey. ; Rabbit didn't give every cat any coffee. ; Lion didn't give every dog any water. ; Elephant didn't give every monkey any soup.

(22) [-**NPI**] **items:** Cat didn't give the giraffes coffee. ; Dog didn't give the bears soup. ; Elephant didn't give the lions honey. ; Rabbit didn't give the monkeys water. ; Elephant didn't give every giraffe tea. ; Bear didn't give every lion ice cream. ; Giraffe didn't give every bear honey. ; Rabbit didn't give every cat coffee. ; Every bear who didn't have ice cream ate chocolate. ; Exactly two giraffes had ice cream. ; Giraffe had ice cream. ; Every cat who was thirsty drank juice.

Experiments 2 and 5: Picture Selection Task[18]

(23) **Example items:** Lion colored the circle. ; Monkey opened the window. ; Monkey colored the squares blue.

(24) **Target items:** Monkey didn't give every rabbit juice. ; Cat didn't give every dog pie. ; Elephant didn't give every giraffe tea. ; Bear didn't give every lion ice cream. ; Giraffe didn't give every bear honey. ; Rabbit didn't give every cat coffee. ; Lion didn't give every dog water. ; Elephant didn't give every monkey soup.

(25) **Controls:** Rabbit didn't give every cat milk. ; Dog didn't give every elephant cheese. ; Monkey didn't give every rabbit ice cream. ; Bear didn't give every monkey cake. ; Lion gave every bear honey. ; Dog gave every elephant cheese. ; Elephant gave every dog cake. ; Elephant gave every giraffe juice. ; Cat didn't give Dog every apple. ; Giraffe didn't give Lion every strawberry. ; Elephant didn't give Giraffe every flower. ; Rabbit didn't give Cat every cupcake. ; Elephant didn't give Giraffe juice. ; Rabbit didn't give Cat tea. ; Lion didn't give Bear honey. ; Dog didn't give Elephant ice cream. ; Giraffe got every flower. ; Rabbit got every strawberry. ; Bear got juice. ; Elephant got tea.

Experiments 2, 3, and 4: Acceptability Judgment Task

(26) **Example items:** Every monkey who picked any fruit had a basket. ; Every cat who had a basket picked any fruit. ; Exactly two cats picked any fruit.

(27) **Target items:** Monkey didn't give every rabbit any juice. ; Cat didn't give every dog any pie. ; Elephant didn't give every giraffe any tea. ; Bear didn't give every lion any ice cream. ; Giraffe didn't give every bear any honey. ; Rabbit didn't give every cat any coffee. ; Lion didn't give every dog any water. ; Elephant didn't give every monkey any soup.

(28) **Controls:** Giraffe had any ice cream. ; Cat had any cheese. ; Lion drank any coffee. ; Elephant ate any cake. ; Pig didn't have any honey. ; Monkey didn't pick any fruit. ; Rabbit didn't drink any tea. ; Dog didn't eat any pie.

Experiments 3 and 4: Covered Picture Task and Implicature Task

The example and target items were the same as in Experiment 2; the only difference was in the presence/absence of certain control items. Below are all the control items from the Covered Picture Task of Experiment 3 and the implicature task of Experiment 4.

(29) **Controls:** Bear didn't give every monkey honey. ; Giraffe didn't give every lion chocolate. ; Rabbit didn't give every cat milk. ; Dog didn't give every elephant cheese. ; Monkey didn't give every rabbit ice cream. ; Bear didn't give every monkey cake. ; Lion gave every bear honey. ; Dog gave every elephant cheese. ; Elephant gave every dog cake. ; Elephant gave every giraffe juice. ; Cat didn't give Dog every apple. ; Giraffe didn't give Lion every strawberry. ; Elephant didn't give Giraffe every flower. ; Rabbit didn't give Cat every cupcake. ; Elephant didn't give Giraffe juice. ; Rabbit didn't give Cat tea. ; Lion didn't give Bear honey. ; Dog didn't give Elephant ice cream.

[18] The stimuli in Experiment 5 included a modification to the target items administered to the intervention group, as explained in the relevant sections.

6.8.3 Experiment 5: Intervention group questionnaire

Questionnaire

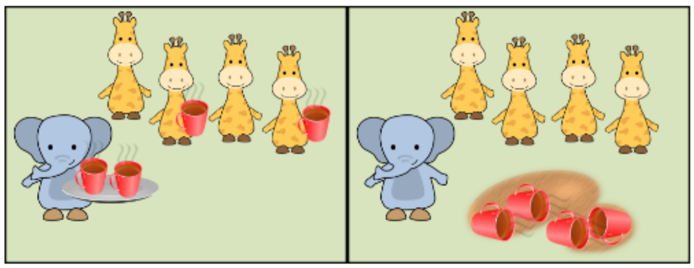
In this task you saw sentences like:

Elephant didn't give every giraffe any tea.

Question 1: How acceptable did you find the word 'any' in the sentence, with 1 being completely odd, and 5 being completely acceptable?

Completely odd ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Completely acceptable*

Question 2: This sentence was paired with the following pictures:



Picture 1 | Both pictures | Picture 2

Which answer did you select?

○ Picture 1 ○ Picture 2 ○ Both pictures

Question 3: Please explain how you decided on your answer. For example, if you chose Picture 1, tell us why you chose Picture 1.

Figure 6.18: Final questionnaire in the repair strategy experiment

Questionnaire Q3: Justifications for picture choices

'Picture 1' selections

1. "if elephant didn't give 'every' giraffe tea then he gave 'some' of the giraffes tea"
2. "I chose it because he wanted to give some but didn't give it to everyone."
3. "I believe I selected Picture 1. I thought that it possibly meant not every giraffe got tea, so that meant some may have."
4. "Implicit with every is that some did get tea, I think."
5. "Because it means he gave some tea but not every one of them."
6. "I thought the use of every implied that some got tea."
7. "because its some without tea"
8. "Because it said he didn't give every giraffe tea , it seemed that he did give it to at least one."
9. "Because the sentence leads you to believe he gave some tea but not all"

'Both pictures' selections

1. “In both pictures not every giraffe has tea, so both are correct.”
2. “Not all giraffes received tea in both pictures.”
3. “Chose both pictures because not every giraffe has tea in both pictures.”
4. “I chose both pictures because in both pictures either no giraffes got tea or just half of them got tea.”
5. “Neither picture showed all giraffes with tea”
6. “The sentence is a bit ambiguous and could either mean that the elephant didn’t give any tea to thegiraffes, or that he didn’t give tea to all the giraffes.”
7. “I chose both because in both pictures, not every giraffe has tea”
8. “In both pictures not every giraffe has tea”
9. “to me it was that some of them or none of them got tea just not all of them”
10. “None of the giraffes have tea.”

‘Picture 2’ selections

1. “No giraffe had any tea at all, while in Picture 1 SOME giraffes did have some tea”
2. “not every giraffe got tea”
3. “Not every giraffe had tea.”
4. “not all the giraffes have tea in both pictures”
5. “I chose picture 2 because every giraffe didn’t get any tea. I would have chose Picture 1 if they used some instead of every giraffe.”
6. “I just went with my gut, which in this case was telling me picture 2 (if I remember correctly)”
7. “It say it didn’t give EVERY giraffe any tea.”

Acknowledgements We would like to thank Maria Aloni, Amir Anvari, Gennaro Chierchia, Luka Crnič, Anna Czypionka, Jakub Dotlačil, Manuel Križ, Jeremy Kuhn, Mora Maldonado, Salvador Mascarenhas, Rick Nouwen, Philippe Schlenker, and Benjamin Spector for helpful discussion, critiques, and questions, and Dorothy Ahn for allowing us to use her artwork in our experiments. We are also grateful to three anonymous reviewers for their helpful comments, questions, and suggestions. The research leading to this work was supported by the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n.313610, by ANR-10-IDEX-0001-02 PSL* and ANR-10-LABX-0087 IEC, and by the Australian Research Council Centre of Excellence in Cognition and its Disorders (CE110001021).

Part III

Implicatures

Chapter 7

Blind ignorance inferences of embedded disjunction

This chapter is a reproduction of the following manuscript:

Denić, Milica. Blind ignorance inferences of embedded disjunction. Manuscript.

Note: This manuscript is to be submitted to a journal. An earlier version of a part of it was published as:
Denić, Milica. 2018. A new case of pragmatically deviant embedded disjunction. In *Proceedings of SALT 28*, pp.454–473.

Abstract

Utterances such as 'Everyone took an apple or a pear' trigger inferences of the form 'Someone took an apple and someone took a pear' (Klinedinst, 2007; Crnič et al., 2015, a.o.). Based on a novel observation that the derivation of these inferences is sensitive to the number of people in the domain, as well as to the number of disjuncts in a sentence, we propose that scalar implicatures are not derived by negating a set of logically stronger alternatives to the utterance, as is standardly assumed, but a set of alternatives that are sufficiently informative given the literal meaning of the utterance, with the measure of informativeness being that of likelihood.

7.1 Introduction

This paper addresses the following question: is scalar implicature computation sensitive to the logical relation of entailment between alternatives, or is it sensitive to probabilistic informativeness? This question has been very recently discussed by Fox and Katzir (2019), who demonstrate that the existing arguments in the literature in favor of the latter do not stand up to closer scrutiny. Fox and Katzir (2019) thus conclude that probabilities either do not play any role in implicature derivation, or if they do, they do so in a modular way, i.e. dissociated from contextual knowledge.

In this paper, we will discuss a family of novel observations, which we use to argue in favor of the idea that probabilistic informativeness rather than logical entailment is at heart of implicature derivation. We will also see evidence that this probabilistic informativeness computation cannot have access to contextual knowledge (in other words, it has to be computed in a modular way). This is in line with the discussion in Fox and Katzir (2019).

The first observation we will discuss is the effect of domain size on inferences of embedded disjunctions. Consider (1a) and (1b). These two sentences are structurally nearly identical: they differ essentially in the size of the domain of the universal quantifier. Strikingly, however, (1a) and (1b) trigger very different inferences: as we will discuss in detail in Section 7.2, while (1a) is interpreted preferably with distributive inferences, (1b) is interpreted preferably with ignorance inferences.

(1) a. All 20 of Mary's friends are French or Spanish.
    b. Both of Mary's friends are French or Spanish.

The second observation we will discuss is why the sentence in (2a) and other structurally similar sentences are deviant. The meaning (2a) should be able to convey is (2b); yet clearly it cannot be used to do so.

(2) a. #Each of these three girls is Mary, Susan, or Jane.
    b. These girls are Mary, Susan, and Jane.

Based on these two observations, we will argue that (i) probabilistic informativeness plays a role in implicature derivation, and (ii) that this probabilistic informativeness is computed in a modular way.

Part 1: Ignorance and distributive inferences of embedded disjunction

In the first part of the paper, we will focus on the first observation: the different inferential pattern of the sentences in (1). We will first go over the empirical generalizations the account has to capture. We will then introduce the exhaustification approach to implicature derivation (Fox, 2007; Chierchia et al., 2012), and show that it fails to capture this difference. The reason why it fails to capture the difference is simple: the entailment relations between (1a) and its alternatives are completely parallel to the entailment relations between (1b) and its alternatives. The failure to capture the different inferential patterns of the sentences in (1) is however by no means unique to the exhaustification approach: any approach according to which the implicature derivation is sensitive to the entailment relations between a sentence and its alternatives will face the same problem (for instance, standard neo-Gricean approaches such as Sauerland (2004)).

7.2 Empirical generalizations

Non-quantified sentences with disjunction such as (3a) trigger *ignorance implicatures* about their disjuncts (Gazdar 1979, a.o.): for instance, (3) typically implies (3a,b).

(3) John is French or Spanish.
    a. $\rightsquigarrow$ The speaker is ignorant about whether or not John is French.
    b. $\rightsquigarrow$ The speaker is ignorant about whether or not John is Spanish.

When the disjunction is in the scope of a quantifier, however, as in (4), ignorance implicatures are usually absent: even though the speaker can in principle both believe (4) and be in the epistemic state is as in (4b), we do not typically infer (4b) upon hearing (4). In fact, (4) typically triggers the so-called *distributive inferences* in (4a) (Chemla (2009); Chemla and Spector (2011); Crnič et al. (2015); Chierchia et al. (2012); Fox (2007); Klinedinst (2007); Spector (2006), a.o).

(4) **All 20 of Mary's friends are French or Spanish.**

a. ⇝ At least one of them is French.
⇝ At least one of them is Spanish.
b. ⇝̸ The speaker is ignorant about whether or not all 20 of them are French.
⇝̸ The speaker is ignorant about whether or not all 20 of them are Spanish.

Strikingly, this inference pattern is sensitive to the cardinality of the restrictor of the quantifier. Consider the sentence (5). (5) is minimally different from (4) in that the cardinality of the restrictor of the universal quantifier is 20 in (4) and two in (5). This change in cardinality reverses the inference pattern of (5) as compared to (4): (5) no longer triggers the distributive implicatures in (5a). Instead, it suggests that the speaker is ignorant about whether both of Mary's friends are French and whether both are Spanish, as in (5b).

(5) **Both of Mary's friends are French or Spanish.**
a. ⇝̸ At least one is French.
⇝̸ At least one is Spanish.
b. ⇝ The speaker is ignorant about whether or not both are Spanish.
⇝ The speaker is ignorant about whether or not both are French.

The phenomenon thus seems to be that when the cardinality of the restrictor in a quantified sentence with embedded disjunction is large, the sentence is most naturally interpreted with distributive and not with ignorance inferences; when the cardinality of the restrictor is small, the sentence is most naturally interpreted with ignorance inferences and not with distributive inferences.

What happens when the cardinality of the restrictor takes values intermediate to two and 20? As far as the linguistic intuitions go, there doesn't seem to be a clear threshold n such that when the cardinality of the restrictor is smaller than n, the sentence triggers ignorance inferences, and when it is equal or larger, the sentence triggers distributive inferences. Instead, the effect appears to be ***gradient***: (6a) is still reported to be more naturally interpreted with ignorance and not distributive inferences, (6b) less clearly so, etc.

(6) a. All three of Mary's friends are French or Spanish.
b. All four of Mary's friends are French or Spanish.
c. All five of Mary's friends are French or Spanish.
...
d. All twenty of Mary's friends are French or Spanish.

At first glance, it appears that the larger the cardinality of the restrictor, the more distributive inferences are derived and not ignorance inferences. However, the empirical picture is slightly more complicated. Consider (7) and (8). The restrictor of the universal quantifier has the same cardinality in these two examples (four), but the number of disjuncts is different: there are two disjuncts in (7), and four disjuncts in (8). (7) is reported to be more easily intepreted with distributive inferences and without ignorance inferences than (8).

(7) All four of Mary's friends are French or Spanish.

(8) All four of Mary's friends are French, Spanish, German, or Portuguese.

A more precise description of the data is thus that the inference pattern of sentences such as (3) is sensitive to ***the ratio between the cardinality of the restrictor and the number of disjuncts***, rather than just to the cardinality of the restrictor.

Finally, the data discussed so far consists of universally quantified sentences. However, the observed inference pattern appears to hold more generally with other forms of quantification. For instance, (9) is more easily interpreted with distributive inferences and not with ignorance inferences than (10).

(9) Twenty of Mary's friends are French or Spanish.

(10) Two of Mary's friends are French or Spanish.

Let us summarize the empirical picture so far: what determines whether a quantified sentence with an embedded disjunction triggers ignorance inferences or distributive inferences? Note that under certain assumptions, (4) carries more information (it is more surprising) than (5). This is true in particular under the following assumptions. Suppose that A = 'French or Spanish' is a predicate such that for any two relevant individuals, the probability of them satisfying A are independent from one another and equal to $p_A \in]0,1[$. For a domain size $n \in \mathbb{N}$, the probability that everyone is French or Spanish is then p_A^n, and this number decreases as n increases. It is also easy to show that under similar assumptions (7) is more informative than (8), and (9) than (10). The empirical generalization that emerges from this is in (11).

(11) ***Empirical generalization:*** The more informative the literal meaning of a quantified sentence with an embedded disjunction, the more likely we are to derive distributive and not ignorance inferences.

7.3 Pragmatic and semantic approaches to implicatures

We have introduced two types of implicatures so far: distributive and ignorance implicatures. Distributive inferences are usually assumed to be a type of *scalar implicatures*. Let us see how scalar and ignorance implicatures are derived according to the exhaustification approach.

Let us first consider a simple case such as (3), repeated in (12). We have already seen that (12) triggers ignorance inferences in (12a). In addition to them, (12) also triggers the scalar implicature in (12b).

(12) John is French or Spanish.

a. The speaker is ignorant about whether John is French (Spanish).
b. John isn't French and Spanish.

According to the exhaustification approach to implicatures, scalar implicatures are not the result of pragmatic reasoning. They are assumed to be part of the logical meaning of the sentence as a result of the semantics of a silent exhaustivity operator *exh*. This operator is assumed to be present in the logical form of a sentence, as in (13).

(13) [*exh* [John is French or Spanish]]

The semantics of this silent operator *exh* is given in (14), it is very similar to that of the focus operator *only* (Chierchia, 2006; Fox, 2007; Chierchia et al., 2012). In short, the semantic import of *exh* when it attaches to a sentence S is to negate as many alternatives as possible from the alternative set $ALT(S)$. There is however one restriction on the alternatives which can be negated: only those alternatives which are *innocently excludable* can be negated. Innocently excludable (IE) alternatives are those which appear in every maximal set of alternatives which can be negated consistently with the original utterance (cf. (14b)).

(14) a. $Exh(p) = p \wedge \bigwedge_{q \in IE(p)} \neg q$
b. $IE(p) = \bigcap\{A' \subseteq ALT(p)$: A' is a maximal set in $ALT(p)$ which can be negated consistently with $p\}$

What alternatives does a sentence activate? Simplifying somewhat, the formal alternatives (FA) a sentence S activates are standardly assumed to be obtained by replacing the constituents of S with another expression of the same type and smaller or equal structural complexity (Katzir, 2007; Fox and Katzir, 2011). The final set of alternatives a sentence S activates $ALT(S)$ in a given context are all those formal alternatives which are relevant in that context (cf. (15)).

(15) **Alternatives of a sentence *S***:
$ALT(S) = \{[\![X]\!] : X \in FA(S)\} \cap \{p : p$ is a contextually relevant proposition$\}$

Let us now see how this derives the inferences of (12). We assume that $ALT((12))$ is in (16).

(16) Relevant formal alternatives of (12):
a. John is French.
b. John is Spanish.
c. John is French and Spanish.

There are two maximal sets of alternatives of (12), (17a) and (17b), which can be negated consistently with (12).

(17) a. {John is French, John is French and Spanish}
b. {John is Spanish, John is French and Spanish}

The only innocently excludable alternative of (12) is thus *John is French and Spanish*, as it is the only alternative which appears in both (17a) and (17b). (12), parsed as (13) thus ends up being interpreted as in (18).

(18) John is French or Spanish and he isn't French and Spanish.

How about the derivation of ignorance inferences?

One approach[1] to ignorance inferences is pragmatic in nature: ignorance inferences of unembedded disjunction are a consequence of the maxim of quantity, according to which the speaker should convey all of the relevant information he has. We will adopt the version of the

[1]Another approach to ignorance implicatures within the exhaustification framework derives ignorance implicatures as semantic inferences (Meyer, 2013, 2014; Buccola and Haida, 2018). In other words, ignorance implicatures, just like scalar implicatures, end up being part of the logical meaning of the sentence. In the first part of the paper, we will assume the pragmatic approach to the derivation of ignorance inferences, as it requires less technical machinery. We will introduce the grammatical approach to ignorance inferences in the second part of the paper, as it will become more relevant then.

maxim of quantity in (19), adapted from Fox (2007):

(19) *Basic maxim of quantity:* If two sentences S_1 and S_2 are both relevant to the topic of conversation, and S_1 is more informative than S_2, if the speaker believes both S_1 and S_2 to be true, the speaker should say S_1 rather than S_2.

Let us assume that a sentence S is more informative than another sentence S' iff S (asymmetrically) entails S'. Assume further that in a context in which the sentence (12) is uttered and relevant, the sentences S_1 = *John is French* and S_2 = *John is Spanish* are also relevant. Note that S_1 and S_2 are more informative than (12) (both S_1 and S_2 entail (12)): the maxim of quantity thus licenses the inferences in (21).

(20) The speaker doesn't believe that John is French (Spanish)

Assuming that the speaker believes his own utterance (12) (maxim of quality), the inferences in (20) amount to ignorance inferences in (21):

(21) The speaker doesn't know whether or not John is French (Spanish)

More generally, assuming together with Fox (2007) that relevance is closed under conjunction and negation, ignorance inferences are predicted to be derived about any relevant sentence S whose truth is not settled by the utterance U. The reason is that if U is relevant and S is relevant, so is $U \wedge S$, as well as $U \wedge \neg S$. As both of these are more informative than U, the maxim of quantity licenses inferences that the speaker doesn't believe $U \wedge S$ or $U \wedge \neg S$: together with the maxim of quality this amounts to the ignorance inference about S.

7.3.1 The case of embedded disjunction

Let us now see what implicatures the two approaches predict for (4) and for (5), repeated here.

(4) All 20 of Mary's friends are French or Spanish.

(5) Both of Mary's friends are French or Spanish.

The predictions of any theory of implicatures for a given sentence depend on the alternatives that the sentence is assumed to activate. Sentences in (4) and (5) have two scalar items, both of which can activate alternatives: the universal quantifier (*all*, *both*), and the disjunction. If both of these scalar items activate their alternatives, the set of formal alternatives consists of all sentences in which the universal quantifier, the disjunction, or both, are replaced by alternative expressions they activate. Concretely, for (4) and (5), this means that the alternatives are as in (22) and (23) respectively: we will henceforth refer to the set of alternatives in (22) and (23) as *ALT(all-or)*.

(22) Alternatives of (4): attempt 1
  a. All 20 are French
  b. All 20 are Spanish
  c. All 20 are French and Spanish
  d. Some are French
  e. Some are Spanish
  f. Some are French or Spanish
  g. Some are French and Spanish

(23) Alternatives of (5): attempt 1
  a. Both are French
  b. Both are Spanish
  c. Both are French and Spanish
  d. Some are French
  e. Some are Spanish
  f. Some are French or Spanish
  g. Some are French and Spanish

Another possibility is that only one of the two scalar items activates alternatives. If only the disjunction activates its alternatives, the formal alternatives of (4) and (5) are in (24) and (25) respectively.[2] We will henceforth refer to the set of alternatives in (24) and (25) as *ALT(or)*.[3]

(24) Alternatives of (4): attempt 2
  a. All 20 are French
  b. All 20 are Spanish
  c. All 20 are French and Spanish

(25) Alternatives of (5): attempt 2
  a. Both are French
  b. Both are Spanish
  c. Both are French and Spanish

Let us now see which implicatures are predicted for (4) and for (5) by the exhaustification approach, under each of the two sets of alternatives.[4] To give a preview of what follows,we will see that ignorance inferences are predicted for both (4) and (5) under the set of alternatives *ALT(all-or)*, while distributive inferences are predicted for both (4) and (5) under the set of alternatives *ALT(or)*.

If (4) activates the alternatives *ALT(all-or)*, the maximal sets of alternatives which can be negated consistently with (4) are in (26):

(26)
  a. {All 20 are French, All 20 are Spanish, Some are French and Spanish, All 20 are French and Spanish}
  b. {All 20 are French, Some are French, Some are French and Spanish, All 20 are French and Spanish}
  c. {All 20 are Spanish, Some are Spanish, Some are French and Spanish, All 20 are French and Spanish}

This means that the only innocently excludable alternatives will be (22c) and (22g) (they are the only alternatives which appear in all three sets in (26)). The predicted scalar implicatures will thus be the negations of (22c) and (22g). Assuming that the sentences S_1 = *Everyone is French*, S_2 = *Everyone is Spanish*, S_3 = *Someone is French*, S_4 = *Someone is Spanish* are relevant, ignorance inferences about them are derived as a consequence of the maxim of quantity in (19).

How do these predictions match the actual inferences people get with (4) and for (5)? We have seen that (5) triggers ignorance inferences while (4) does not (or at least to a very differ-

[2]Note that if only the universal quantifier activates its alternatives, the only alternative of (4) and (5) would be 'Some is French or Spanish': if the restrictor of the universal quantifier is non-empty, this alternative is entailed bt the original sentence, so no implicatures are derived.

[3]Fox (2007) and Magri (2009b) assume that the *ALT(or)* alternatives are the only alternatives that sentences such as (4) and (5) activate; see also the discussion in Bar-Lev and Fox (2017), fn. 7.

[4]Note that the predictions are the same under certain neo-Gricean approaches to implicatures, for instance, that of Sauerland (2004).

ent degree). Assuming that sentences (4) and (5) activate the alternatives *ALT(all-or)*, correct inferences are predicted for (5) but not for (4).

Let us now see what the predictions are under the set of alternatives *ALT(or)*.

All the alternatives in *ALT(or)* can be negated consistently with (22), therefore all of the alternatives in *ALT(or)* are innocently excludable. This accounts for the distributive inferences: again, the negation of (24a) together with (4) entails that some of Mary's friends are Spanish, and the negation of (24b) together with (4) entails that some of Mary's friends are French. The problem is, of course, that if both (4) and (5) activate the alternatives in *ALT(or)*, both of them are now predicted to trigger distributive inferences.

To summarize, (4) gives rise to distributive inferences, while (5) gives rise to ignorance inferences. The exhaustifiction approach to implicature derivation, as it stands, cannot capture this difference. It is worthwhile stressing again that the reason for this is fully general: what is heart of the exhaustification approach (and many other approaches to implicature derivation) is the logical relation between a sentence and its alternatives; crucially, to the extent that (4) and (5) activate parallel sets of alternatives, they stand in the same logical relations to their alternatives, and will thus necessarily be predicted to have the same implicatures. In particular, under the exhaustification approach, depending on what set of alternatives is chosen, either ignorance inferences are predicted for both sentences (for the alternatives in *ALT(all-or)*), or distributive inferences are predicted for both sentences (for the alternatives in *ALT(or)*).

We will now turn to a proposal for how to constrain implicature derivation so that the cardinality of the restrictor effect is captured.

7.4 Proposal: constraint on alternative pruning

7.4.1 Informativeness and alternative pruning

The main piece of data we want to account for is the difference in inferences which (4) and (5) trigger. As a reminder, (4) is interpreted preferably with distributive and not with ignorance inferences, while the opposite is the case for (5).

(4) All 20 of Mary's friends are French or Spanish.

(5) Both of Mary's friends are French or Spanish.

Two assumptions will be central to our proposal. The first assumption is that some of the formal alternatives can sometimes not be taken into account when implicatures are calculated — the term that is usually employed for this is *alternative pruning* (Fox and Katzir, 2011; Katzir, 2014; Crnič et al., 2015). We have in fact already introduced the notion of pruning by defining the set of alternatives that are considered for implicature derivation as the intersection of formal alternatives with contextually relevant propositions as in (15), repeated here.

(15) **Alternatives of a sentence *S***:
$ALT(S) = \{[\![X]\!] : X \in FA(S)\} \cap \{Y : Y \text{ is a contextually relevant proposition}\}$

Which alternatives are pruned is thus typically assumed to depend on contextual relevance considerations. For instance, while (27) typically triggers the scalar implicature in (27a), in some contexts it can also trigger the stronger implicature in (27b) or (27c) (Horn, 1972; Katzir, 2014).

(27) John did some of the homework.

a. John didn't do all of the homework.
b. John didn't do most of the homework.
c. John didn't do much of the homework.

The fact that (27) doesn't always trigger the implicatures in (27b) and (27c) is usually taken as an argument that the formal alternatives of (27) 'John did most of the homework' and 'John did much of the homework' are sometimes not part of *ALT((27))* (i.e. they are sometimes pruned from *ALT((27))*) for contextual reasons).

The second assumption is that sentences such as (4) and (5) activate *ALT(all-or)* alternative set, and that the alternative set *ALT(or)* is derivable from *ALT(all-or)* by pruning some of the alternatives from *ALT(or)*.

The proposal we will put forward to account for the contrast between (4) and (5) is that alternative pruning is, in addition to contextual relevance, also sensitive to how informative alternatives are, with a measure of informativeness to be made precise in what follows. In other words, we propose that the alternative set considered for implicature derivation of a sentence *S* is as in (28):

(28) **Alternatives of a sentence *S*: proposal**
$ALT(S) = \{[\![X]\!] : X \in FA(S)\} \cap \{Y : Y \text{ is a contextually relevant proposition}\} \cap$
$\{Z : Z \text{ is an informative proposition}\}$

We propose that the measure of informativeness employed for pruning is that of likelihood: the more unlikely the propositional meaning of an alternative sentence is, the more informative the alternative sentence is. For presentation purposes, we start with a very simple version of the proposal which only cares about the informativeness of the alternatives, and not about about the informativeness of the original utterance.

(29) **Informative propositions and pruning: proposal *(to be revised)***
Let A be a formal alternative of S, and $P(A)$ the likelihood of its propositional meaning. The larger *P(A)* (and thus the less informative *A*!), the more likely we are to prune *A* from *ALT(S)*.

An intuitive reason for why something like (29) might hold of pruning is that the more likely an alternative *A* is to be true, the less pressure there is for the speaker to utter *A*, and thus the less pressure there is to consider it as an alternative utterance the speaker could have said instead of his original utterance *S*. This idea makes the proposal conceptually related to rational speech act (RSA) models (Bergen et al., 2016; Goodman and Stuhlmüller, 2013, a.o.), in which the notion of probabilistic informativeness of alternatives plays precisely this role in implicature derivation.[5]

Let us see how this proposal captures the difference between (4) and (5). In particular, for some domain size *n*, for a sentence of the form (30), let us consider what happens with its alternatives (30a,b).

[5]Benjamin Spector (p.c.) notes that despite this intuitive connection between the current proposal and the RSA approach, the observed empirical pattern of distributive and ignorance inferences is not straightforwardly captured by the RSA models.

(30) All of n people are A or B.

a. Some of n people are A, Some of n people are B

b. All of n people are A, All of n people are B

We submit that it is reasonable to assume that **the larger the domain of people is, the more likely it is that there is someone among them who is A**. This is true in particular under the following assumptions (we will return later to the question of whether these assumptions are justified): (1) Suppose that A, B are predicates such that for any relevant individual the probability that they satisfy A equals the probability that they satisfy B equals $p \in]0,1[$ (in other words, assume that the probability distribution over predicates is uniform). (2) Assume further that whether or not an individual satisfies A is independent of whether this individual satisfies B, and of which predicates other individuals satisfy.

Fact 1. For a domain size $n \in \mathbb{N}$ then, $P(\text{Some of the } n \text{ individuals are } A) = 1-(1-p)^n$ (i.e. 1 - P(All n individuals are not in A)). Crucially, this probability increases with the domain size n (see Fact 1 in Appendix A).

Fact 2. For any domain size $n > 1$, it is more likely that someone in that domain is A than that everyone in that domain is A. This is because, with the above notations, $1-(1-p)^n > p^n$ (see Fact 2 in Appendix A).

We illustrate this the effect of the domain size n on the informativeness of a subset of alternatives of (30) in Table 7.1.

A (alternative)	n	p	$P(A)$
All n are A	6	0.2	$0.2^6 = 0.00006$
Some n are A	6	0.2	$1-(1-0.2)^6 = 0.74$
All n are A	2	0.2	$0.2^2 = 0.04$
Some n are A	2	0.2	$1-(1-0.2)^2 = 0.36$

Table 7.1: $P(A)$ with $p = 0.2$

From Fact 1 above, we may conclude that alternatives of the form 'Some of the n individuals are A' are more likely for larger n, and therefore more likely to be pruned for larger ns. And this is true in principle for A=French, A=French or Spanish, A=French and Spanish, etc. It thus follows that the set of alternatives that we will end up with is more likely to be the restricted set for larger ns, the one yielding distributive inferences, and the full set for smaller ns, the one yielding ignorance inferences. Fact 2 above suggests that these 'all' alternatives will survive longer than the 'some' alternatives, so if anything is to stay, it will be the alternative 'All of the n individuals are A', rather than 'Some of the n individuals are A'.

Concretely, this means that we are more likely to end up with the set of alternatives such as in Table 7.2 for (4) than with a parallel set of alternatives for (5). As we have seen in Section 7.3.1, with this set of alternatives, distributive inferences are derived.

Furthermore, we are more likely to end up with the set of alternatives as in Table 7.3 for (5) than with a parallel set of alternatives for (4): we have seen in Section 7.3.1 that with this set of alternatives, ignorance inferences are derived.[6]

[6]In table cells of Table 7.4 and 7.5 '...' in the 'Alternatives' column is intended to convey 'and other alterna-

Alternatives of (4)	Inferences
All 20 are French	**Some are Spanish**
All 20 are Spanish	**Some are French**
All 20 are French and Spanish	Not all are French and Spanish
~~**Some are French**~~	—
~~**Some are Spanish**~~	—
~~Some are French or Spanish~~	(entailed by (5))
Some are French and Spanish	No one is French and Spanish
All 20 are French and Spanish	Not all are French and Spanish

Table 7.2: Inferences of (4) — the least informative alternatives pruned

Alternatives of (5)	Inferences
Both are French	**ignorance**
Both are Spanish	**ignorance**
Both are French and Spanish	Not both are French and Spanish
Some are French	**ignorance**
Some are Spanish	**ignorance**
Some are French or Spanish	(entailed by (5))
Some are French and Spanish	No one is French and Spanish
Both are French and Spanish	Not both are French and Spanish

Table 7.3: Inferences of (5) — no alternatives pruned

This proposal (29) can thus capture that the smaller the cardinality of the restrictor, the more ignorance inferences are derived, and the less distributive inferences are derived.

7.4.2 Refinement of the proposal

The proposal in (29) has to be refined somewhat. There are two empirical reasons to do so. The first is that we cannot yet capture that this effect is not restricted to universally quantified sentences. As a reminder, consider (9) and (10), repeated here. The situation is similar to what we have seen with the universally quantified sentences: (9) is most naturally interpreted with distributive inferences, and (10) with ignorance inferences.

(9) Twenty of Mary's friends are French or Spanish.

(10) Two of Mary's friends are French or Spanish.

Let us see why the proposal in (29) cannot capture this even if one assumes that a sentence such as (9) activates the alternatives in (31) and a sentence such as (10) activates the alternatives in (32), and in addition to that allows for an informativeness-based pruning as stated in the proposal in (29).

tives with the same number of disjuncts or conjuncts'; '...' in the 'Inferences' column is intended to convey 'the inferences derived from the alternatives with the same number of disjuncts/conjuncts'.

(31) Alternatives of (9)

a. Twenty are French
b. Twenty are Spanish
c. Some are French
d. Some are Spanish

(32) Alternatives of (10)

a. Two are French
b. Two are Spanish
c. Some are French
d. Some are Spanish

The reason why we cannot capture this is that *P(Some of Mary's friends are French)* is solely dependent on the number of Mary's friends, and the number of her friends is independent of how many of them are French or Spanish. In other words, we will be equally likely to prune the existential replacement alternatives from the (31) as from (32). In other words, the prediction of (29) is that we should be equally likely to derive ignorance inferences for (9) and for (10) (likewise for distributive inferences), contrary to the facts.

The second reason to refine the account is that we cannot capture yet that the inference pattern is sensitive to the ratio between the cardinality of the restrictor and the number of disjuncts, rather than just to the cardinality of the restrictor. As a reminder, compare the sentences (7) and (8), repeated below: (7) is reported to be more easily interpreted with distributive and without ignorance inferences than (8).

(7) All four of Mary's friends are French or Spanish.

(8) All four of Mary's friends are French, Spanish, German, or Portuguese.

Let us see why the proposal in (29) cannot capture this by focusing on inferences predicted for (8). If (8) activates the alternatives as in Tables 7.4 and 7.5, it is predicted to trigger ignorance inferences when no alternatives are pruned as in Table 7.4, and distributive inferences when the alternatives of the form 'Some are French', 'Some are French or Spanish' etc. are pruned, as in Table 7.5.

Alternatives of (8)	Inferences
All 4 are French...	**ignorance**
All 4 are French or Spanish...	**ignorance**
All 4 are French or Spanish or German...	**ignorance**
All 4 are French and Spanish...	Not all 4 are French and Spanish...
Some are French...	**ignorance**
Some are French or Spanish...	**ignorance**
Some are French or Spanish or German...	**ignorance**
Some are French and Spanish...	No one is French and Spanish

Table 7.4: Alternatives of (8) — deriving ignorance inferences

Alternatives of (8)	Inferences
All 4 are French...	**Someone is Spanish, German, or Portuguese**...
All 4 are French or Spanish...	**Someone is German or Portuguese**...
All 4 are French or Spanish or German...	**Someone is Portuguese**...
All 4 are French and Spanish...	Not all 4 are French and Spanish
~~**Some are French**...~~	—
~~**Some are French or Spanish**...~~	—
~~**Some are French or Spanish or German**...~~	—
Some are French and Spanish...	No one is French and Spanish

Table 7.5: Alternatives of (8) — deriving distributive inferences

The problem of the proposal in (29) is straightforward: given that pruning was proposed to depend on the likelihood of the propositional meaning of the alternative, alternatives such as 'Someone is French', 'Someone is German' are equally likely to be pruned for (7) and for (8). In other words, the prediction of (29) is that we should be equally likely to derive ignorance inferences for (7) and for (8) (likewise for distributive inferences), contrary to the facts.

A very minor refinement of the proposal, in (33), solves these two problems.

(33) **Informative propositions and pruning: proposal *(final)***
Let A be a formal alternative of S, and $P(A \mid S)$ the conditional likelihood of the propositional meaning of A given S. The larger $P(A \mid S)$ (and thus the less informative A given S!), the more likely we are to prune A from *ALT(S)*.

An intuitive reason for why a constraint on pruning such as (33) might hold is related to what has been said to conceptually motivate (29): the more likely an alternative A is to be true given the utterance S (the closer it is to being entailed by S), the less pressure there is for the speaker to utter both A and S instead of only S if he believes both A and S to be true.

Let us now see that this captures the difference between (7) and (8), as well as that it still predicts the difference between (4) and (5). In particular, for some domain size n, for a sentence of the form (34), let us consider what happens with the alternatives (34a,b,c,d).

(34) All of n people are A_1, A_2, or A_3...
   a. Some of n people are A_1, Someone of n people A_2...
   b. All of n people are A_1, All of n people are A_2...
   c. Some of n people is A_1 or A_2...
   d. All of n people are A_1 or A_2...

We submit that it is reasonable to assume that in a domain D of n people it is less likely that there is someone in D who is A_1 when it is known that everyone in D is A_1, A_2, A_3, or A_4, as compared to how likely it is that there is someone in D who is A_1 when it is known that everyone in D is A_1 or A_2. This is true in particular under the following assumptions. Suppose A_1, A_2, ..., A_k are predicates such that for any relevant individual, the probability that they satisfy A_1 equals the probability that they satisfy A_2, ..., equals $p \in]0,1[$. Assume further that whether or not an individual satisfies some predicate is independent of whether this individual satisfies other predicates, and of which predicates other individuals satisfy. For a domain size $n \in \mathbb{N}$

then, *P(Some of the n individuals is* A_1 *|All n individuals are* A_1 *or* A_2*...or* A_k*)* = $P(\exists i, i \in A_1 \mid \forall i, i \in A_1 \cup A_2 ... \cup A_k)$ is equivalent to $1-(1-\frac{p}{1-(1-p)^k})^n$ (see Fact 3 in Appendix A for derivation). Crucially, (i) this probability increases with the domain size n and (ii) this probability decreases with the number of disjuncts k in the original sentence (see Fact 3 in Appendix A).

We illustrate this effect of domain size n and of the number of disjuncts k on the informativeness on a subset of alternatives of (34) in Table 7.6.

A (alternative)	n	k	p	$P(A \mid S)$
All n are A	6	2	0.2	$(\frac{0.2}{1-(1-0.2)^2})^6 = 0.03$
Some n are A	6	2	0.2	$1-(1-\frac{0.2}{1-(1-0.2)^2})^6 = 0.99$
All n are A	2	2	0.2	$\frac{0.2}{1-(1-0.2)^2}^2 = 0.31$
Some n are A	2	2	0.2	$1-(1-\frac{0.2}{1-(1-0.2)^2})^2 = 0.8$
All n are A	6	4	0.2	$(\frac{0.2}{1-(1-0.2)^4})^6 = 0.001$
Some n are A	6	4	0.2	$1-(1-\frac{0.2}{1-(1-0.2)^4})^6 = 0.92$
All n are A	2	4	0.2	$\frac{0.2}{1-(1-0.2)^4}^2 = 0.11$
Some n are A	2	4	0.2	$1-(1-\frac{0.2}{1-(1-0.2)^4})^2 = 0.56$

Table 7.6: $P(A)$ with $p = 0.2$

Because of this we may conclude that the likelihood of pruning alternatives of the form 'Some of the n individuals are A' increases with the domain size n, and decreases with the number of disjuncts k of the original sentence. And this is true in principle for A=French, A=French or Spanish, A=French and Spanish, etc. It thus follows that the set of alternatives that we will end up with is more likely to be the restricted set (i.e. without the alternatives such as 'Some of the n individuals are A') for larger ns and smaller number of disjuncts.

Concretely, this means that we will still be more likely to prune the alternatives of the form (34a) for (4) than for (5). Furthermore, we will be more likely to prune the alternatives of the form (34a) for (7) than for (8): we will thus be more likely to derive distributive inferences for (7) than for (8), and more likely to derive the inference that the speaker doesn't believe that someone is French, that he doesn't believe that someone is Spanish etc. for (8) than for (7), thus deriving some kind of ignorance inferences. It is however not quite as simple to predict what exactly the content of ignorance inferences will be for sentences such as (34), when the number of disjuncts $k > 2$. The reason is that the informativeness of the alternatives of the form (34c) and (34d) will depend on the precise values of the domain size n, of p and of the number of disjuncts k. As the judgments are subtle, we leave it as an open empirical question how exactly the content of ignorance inferences of a sentence such as (34) varies with n and the number of disjuncts k. The predictions of the proposal at least seems to clearly match the intuitions that distributive inferences are derived more for (7) than for (8), and thus the inferences that the speaker is ignorant about at least some aspects of the situation are more likely with (8) than with (7).

Furthermore, the proposal now extends straightforwardly to the cases in (9) and (10): we have in fact already seen that the bigger the n, the bigger P(Someone is A | n people are A or B) (cf. Fact 1 in Appendix A). This means that we will be more likely to prune the existential replacement alternatives from *ALT((9))* than from *ALT((10))*. As it was the case for universally quantified sentences, pruning the existential replacement alternatives will derive distributive inferences for sentences such as (9) and (10), and not pruning them derives ignorance inferences.

7.5 Interim discussion

To summarize, the proposal in (33) accounts for the two aspects influencing the derivation of ignorance inferences of the disjunction: the cardinality of the restrictor and the number of disjuncts in quantified sentences with embedded disjunction. The proposal also extends straightforwardly to sentences headed with quantifiers other than universal (cf. (9) and (10)). As a reminder, the proposal relies on the assumption that alternatives can be pruned under certain considerations (Horn, 1972; Fox and Katzir, 2011; Katzir, 2014). The core of the proposal is that alternative pruning is sensitive to the informativeness of an alternative conditioned on the original utterance. The proposal is relatively independent of the specifics of the mechanism which derives implicatures: we have demonstrated how it can be implemented with the exhaustification-based framework for implicature derivation, but it is in principle also compatible with neo-Gricean approaches (e.g. Sauerland (2004)).

In addition to this, we would also like to point the interested reader to two other implementation options, in Appendices. More specifically:

- In Appendix B, we discuss two other possibilities for the derivation of ignorance and distributive inferences that become available under the assumption that the exhaustifying operator can apply recursively. We show how the empirical pattern can be accounted for under each of the two possibilities.
- In Appendix C, we demonstrate that the proposal applies without any further assumption even if quantified sentences activate the full range of numerical alternatives.

Finally, we would like to point out that the current proposal may be very close in spirit to the proposal in Chemla and Romoli (2015), which was developed for other purposes. In their framework implicatures of a sentence are eliminated if the informativeness of the implicature is too high. According to our proposal, alternatives are eliminated if the informativeness of the alternative is too low. The two ideas 'covary' in most cases, since the implicature is a consequence of the negation of an alternative.

Importantly, however, pruning an alternative from the whole process of implicature derivation (as in the current proposal) may have radically different effects than eliminating an implicature coming out of the presence of this alternative, depending on which set of alternatives is assumed, and which approach to implicature derivation is adopted. To give a concrete example from the empirical domain explored in this paper, we have seen that under the exhaustification approach to implicature derivation, if one assumes that universally quantified sentences activate existential replacement alternatives, pruning those alternatives allows us to derive distributive inferences, and not pruning derives ignorance inferences. Crucially, however, eliminating ignorance inferences would not straightforwardly lead to the derivation of distributive inferences. This fact differentiates the two proposals.

However, this is not to say that the proposal in Chemla and Romoli (2015) cannot be extended to capture for the data discussed here. In particular, there are free parameters in Chemla and Romoli (2015) to be set up in terms of the set of alternatives assumed and the approach to implicature derivation taken in order to be able to fully compare it to the current proposal. We leave this comparison for future work.

7.6 Informativeness and contextual knowledge

The next question we want to address is where the probabilities that are relevant for informativeness-based pruning come from.

In order to compare alternatives based on their informativeness, we have so far adopted two overly simplifying assumptions: (1) that there is some constant prior probability p over predicates (2) that whether or not an individual satisfies a predicate is independent of what other predicates they satisfy, and what predicates other individuals satisfy. If the probabilities used in informativeness-based pruning come from contextual knowledge of the predicates, these assumptions won't be met in most cases, and we might expect different inferential patterns to emerge.

However, the contextual probabilities do not behave like one might expect them to. To see this, consider (35).

(35) Both of Mary's office-mates are from the US or the UK.

Imagine that Mary works in the US, and that this fact translates into high prior probability that her office-mates are from the US (and hence low prior probability that they are from elsewhere). One thus expects that the alternative 'Some of Mary's office-mates are from the US' will be less informative, and thus more likely to be pruned that the alternative 'Some of Mary's office-mates are from the UK'. If the alternative 'Some of Mary's office-mates are from the US' is pruned but not the alternative 'Some of Mary's office-mates are from the UK', (35) should have the inference that not both of them are from the UK, (so at least one is from the US), and the ignorance inference about whether or not both of Mary's office-mates are from the US.

This interpretation does not seem to be readily available: it thus seems that contextual asymmetries in informativeness do not translate straightforwardly into an effect on the inferences triggered by sentences such as (35).

It thus appears that the mechanism which calculates the informativeness of alternatives does not have access to the contextual knowledge of predicates occurring in these alternatives. This point will be the focus of the discussion in the second part of the paper.

Part 2: Blind ignorance inferences derivation

We now move to the second puzzling observation: why sentence (36) and other structurally similar sentences are deviant.

(36) #Each of these three girls is Mary, Susan, or Jane.

7.7 Empirical generalizations

When disjunctions of definite noun phrases are embedded in the scope of a universal quantifier, the result is sometimes unexpectedly deviant. The deviance depends on the predicate that embeds the disjunction. For instance, when the predicate in question is the identity copula as in (37) or the predicate *to write* in (39), the result is deviant. When the predicate in question is minimally different, as the predicate *to be called* in (38) or the predicate *to read* in (40), the result is acceptable.

(37) (Context: Peter invited three girls to the party.)
#Each of those three girls **is** Mary, Susan, or Jane.

(38) (Context: Peter invited three girls to the party.)
Each of those three girls **is called** Mary, Susan, or Jane.

(39) (Context: Tolstoy, Zola and Rowling are great writers.)
#Each of those three writers **wrote** Anna Karenina, Germinal, or Harry Potter.

(40) (Context: Ann, John, and Bob are great students.)
Each of those three students **read** Anna Karenina, Germinal, or Harry Potter.

To see why the deviance of (37) and of (39) is surprising, note that (37) is contextually equivalent to (41), assuming that it is common knowledge that Mary, Susan, and Jane have to be three different individuals (see (43) for definitions of common knowledge, context set and contextual equivalence). Likewise, (39) is contextually equivalent to (42), assuming that it is common knowledge that for any book there can be exactly one singular or plural individual who wrote it[7]. Yet, surprisingly, (37) cannot be naturally used to do convey the meaning of (41), and neither can (39) to convey the meaning of (42).

(41) One of those three girls is Mary, another one is Susan, and yet another one is Jane.

(42) One of those three writers wrote Anna Karenina, another one wrote Germinal, and yet another one wrote Harry Potter.

(43) **Common knowledge:**
A proposition ϕ is commonly known to a group of individuals if and only if all individuals in the group know that ϕ, all know that all know it, all know that all know that all know it, etc. (Stalnaker, 2002)
Context set:
Context set is the set of possible worlds in which all the propositions that are common knowledge between the interlocutors are true.
Contextual equivalence:
Two sentences A and B are contextually equivalent iff they have the same truth value in all the worlds of the context set.

Note that the deviance observed in (37) and (39) is not specific to *each*: the pattern is the same with *every* and *all*.

The observed deviance is also not limited to universally quantified noun phrases; see for instance (44). However, for the simplicity of exposition, we will focus on the disjunction embedded in the scope of a universal quantifier; the main ideas that will be presented carry over to the cases in (44).

(44) a. #These three girls are Mary, Susan, or Jane.
b. #Three of those girls are Mary, Susan, or Jane.

We have observed the deviance of an embedded disjunction with certain predicates, such as the

[7] In the case of co-authorship there would be exactly one plural individual who wrote the book.

identity copula or *to write*, but not with others, such as *to be called* or *to read*. Which property makes a predicate pattern with one group or the other? We will argue that the essential property that identity copula and *to write* have in common is that when their internal argument is provided in sentences such as (45a) and (45b), given world knowledge, they can only be true of a unique (singular or plural) individual.

In other words, given world knowledge, the predicates *to write* (when its internal argument is from a domain of books) and the identity copula (when its internal argument is from a domain of individuals) have the property in (46), which we will refer to as *left-uniqueness*.

(45) a. That girl is Mary.
b. That writer wrote Anna Karenina.

(46) A predicate P is left-unique iff
$\forall y$ in the relevant domain $\forall x[P(x,y) \Rightarrow \forall z[P(z,y) \Rightarrow z = x]]$

To see that the identity copula (when its internal argument is from a domain of individuals) and *to write* (when its internal argument is from a domain of books) are left-unique but not the predicates *to be called* (when its internal argument is from a domain of names) and *to read* (when its internal argument is from a domain of books), observe that the continuations in (47a) and (47c) sound contradictory, but not in (47b) and in (47d).

(47) a. This girl is my sister Susan. #That other girl is my sister Susan too.
b. This girl is called Susan. That other girl is called Susan too.
c. John wrote this book. #Peter wrote this book too.
d. John read this book. Peter read this book too.

To see that the left-uniqueness is indeed relevant for the phenomenon in question, consider what happens when the internal argument of *to write* is not from a relevant domain for it to be left-unique. For instance, when its internal argument is from a domain of letters of the alphabet, the predicate *to write* is not left-unique, and note that (48), which is structurally similar to (37) and (39), is not deviant (it could perfectly be used in a situation in which, for instance, each of John's three students wrote a number of letters on the board):

(48) Each of John's three students wrote the letter A, the letter D, or the letter K on the board.

7.8 Proposal: blind ignorance inferences

Before moving to the proposal, we will discuss an intuitive hypothesis which we will not pursue for empirical reasons. One could plausibly argue that (37) and (39) activate the alternatives in (49), and that these alternatives are according to some standard better in conveying the meaning of (37) and (39).

(37) #Each of these three girls is Mary, Susan, or Jane.

(39) #Each of these three writers wrote Anna Karenina, Germinal, or Harry Potter.

(49) a. These girls are Mary, Susan, and Jane.
b. These writers wrote Anna Karenina, Germinal, and Harry Potter.

There are however deviant sentences which are plausibly instances of the same phenomenon for which there doesn't seem to be a straightforwardly better way to convey the meaning. Consider for instance a situation in which there are two students, a boy and a girl, and the speaker knows that the girl is either Mary or Jane, and that the boy is either Peter or Jack. (50a) is still not a good way for the speaker to communicate this information. The speaker instead has to say something like (50b). To our knowledge, however, so far there is no theory of alternatives according to which (50b) is an alternative of (50a).

(50) a. #Both of these students are Mary, Susan, Peter or Jack.
b. The girl is Mary or Susan, and the boy is Peter or Jack.

Furthermore, it is not clear why on the same grounds (51a,b) are not deviant because of (51c).

(51) a. Each of my students got an A or a B.
b. All of my students got As or Bs.
c. My students got As and Bs.

We take these two data points to suggest that (37) and (39) are not deviant because the sentences in (49) are somehow better at conveying the meaning of (37) and (39).

Our proposal for the deviance of (37) and (39) will build on the conclusions reached in the first part of the paper. We have seen in the first part of the paper that quantified sentences with embedded disjunction trigger ignorance inferences under certain conditions. Consider now what happens if the deviant (37) and (39) trigger ignorance inferences, which are in fact expected given the ratio between the cardinality of the restrictor and the number of disjuncts in these sentences (cf. empirical generalizations in Section 7.2). These ignorance inferences for (37) and (39) are paraphrased in (52) and (53), respectively.

(52) The speaker is ignorant about whether each of these three girls is Mary... (Susan, Jane), and he is ignorant about whether some of these three girls is Mary... (Susan, Jane).

(53) The speaker is ignorant about whether each of these three writers wrote Anna Karenina, ... (Germinal, Harry Potter), and he is ignorant about whether some of these three writers wrote Anna Karenina,... (Germinal, Harry Potter)

One can immediately see the problem with these ignorance inferences. Because of the left-uniqueness of the identity copula and the predicate *to write*, the sentences like those in (54) contradict common knowledge (i.e. they are contextual contradictions). ***The speaker thus cannot possibly be ignorant about them*** — he must know that they are false.

(54) Each of these three girls is Mary, Each of these three writers wrote Anna Karenina

Furthermore, if we take it to be common knowledge that the speaker should believe their utterances, in the case of (37) and (39), again due to the left-uniqueness of the identity copula and the predicate *to write*, ***the speaker cannot be ignorant about the sentences in (55)*** — he must know that they are true.

(55) Some of these three girls is Mary, Some of these three writers wrote Anna Karenina

The ignorance inferences that sentences (37) and (39) might trigger thus contradict common knowledge. We propose that this is the reason why these sentences are deviant (cf. (56)).

(56) If a sentence *S* triggers *ignorance inferences* which contradict common knoweldge, *S* is deviant.

There is in fact already some empirical data which suggests that ignorance inferences are derived blindly to contextual knowledge with the modified numeral *at least*, and that these ignorance inferences result in the deviance of sentences which triggered them. We provide in (57) an example borrowed from Buccola and Haida (2018); similar empirical observations have been first made by Nouwen (2010).

(57) Context: Ann played a card game in which, given the rules, the final score is always an even number of points. Bob knows this, and reports to Carl:
  a. #Ann scored at least 3 points.
  b. Ann scored at least 4 points.

While we will not go into the the question of how ignorance implicatures are derived for a sentence such as (57a) (see for instance Spector (2019) for a recent review article), a possible reason why (57a) is deviant is because it triggers the ignorance implicature that the speaker is ignorant about whether or not Ann scored exactly 3 points, which contradicts contextual knowledge.

The core assumption of the proposal in (56) is that implicatures generally, and ignorance inferences specifically, are derived *blindly* from common knowledge. The idea that the procedure which derives implicatures is blind to contextual knowledge has been in fact already defended by Magri 2009a for the case of scalar implicatures (see also Degen et al. (2015) for experimental data on unexpected patterns of implicature derivation given contextual knowledge). The pragmatically deviant (58) is an example of the cases that motivated this proposal.

(58) #Some Italians come from a warm country.

Simplifying a lot, according to Magri's proposal, (58) is deviant because the conjunction of (58) and its scalar implicature in (59) contradicts common knowledge.

(59) Not all Italians come from a warm country.

Blind derivation of implicatures is easier to accommodate within theories according to which implicatures are derived in grammar, such as exhaustification-based approaches (Fox, 2007; Chierchia et al., 2012), rather than within neo-Gricean views on ignorance implicatures derivation (Sauerland, 2004). This is so because the calculation of implicatures that contradict common knowledge has to be encapsulated from common knowledge, which is incompatible with the view that implicatures are a result of pragmatic inference integrating common knowledge information. If the proposal (56) is on the right track, even ignorance inferences should be derived in grammar (contra Fox (2007), and in line with grammatical approaches to ignorance implicatures such as (Meyer, 2013, 2014)).

7.9 Back to blind pruning

If Magri's proposal according to which (58) is deviant because of its scalar implicature is correct, and our informativeness-based pruning proposal is on the right track, not only does the derivation of the ignorance inferences has to proceed in a blind manner, **but the mechanism which calculates the informativeness of alternatives must be blind to contextual knowledge**

too. The reason is simply that, given contextual knowledge, P(All Italians come from a warm country | Some Italians come from a warm country) = 1: this means that the alternative 'All Italians come from a warm country' should be pruned due to its lack of informativeness from *ALT*(Some Italians come from a warm country), and that the problematic implicature should not arise. This fact reiterates the observations made in the first part of the paper, according to which contextual probabilities do not seem to influence the inferential pattern of embedded disjunctions in the ways in which one might expect them to (cf. discussion in Section 7.6).

Likewise, informativeness-based pruning that is blind to contextual knowledge is necessary to account for the deviance of (37) within our approach. To see why this is the case, recall that in order to derive ignorance inferences of (37), we needed to not prune the alternatives in (60) from *ALT*((37)).

(60) Someone is Mary,... (Susan, Jane), Someone is Mary or Susan,... (Susan or Jane, Mary or Jane)

However, given common world knowledge P(Some of the girls is Mary| Each of the girls is Mary, Susan, or Jane) = 1, and similarly for all of the alternatives from (60).

This means that, if the proposal in (33) is to capture that alternatives in (60) aren't pruned from the alternative set of (37), the computation of informativeness according to the proposal in (33) has to be blind to (most of) common knowledge: the only things that seem to matter are domain size and logical words (quantifiers, disjunctions etc.) in the sentences. In other words, this means that P(Someone is A | Everyone is A or B or C) is not influenced by the common knowledge about predicates A, B, C.

7.10 Exhaustification approach to ignorance implicatures

The data presented above can be taken as evidence that ignorance inferences are indeed derived in grammar. We thus need *a theory* according to which ignorance inferences are derived in grammar: importantly, this theory should be compatible with the pruning proposal in (33) to capture the inferential pattern of embedded disjunction in quantified sentences discussed in Section 7.2. In this section, we will provide an example of such a theory. We will build on the grammatical theory of ignorance implicatures put forward in Meyer (2013, 2014), which we will modify minimally so that it succeeds in incorporating the pruning proposal in (33) in a relatively straightforward way. More work is needed however to motivate this departure from Meyer's theory independently of the infortmativeness-based pruning proposal, or to find another way to put the pruning proposal together with Meyer's original theory.

Let us first present the main pieces of the grammatical theory of implicatures from Meyer (2013, 2014) before we introduce the aforementioned modification.

(61) **Piece 1:** There is a silent modal operator K in language, with semantics as follows:
$[[K_x\phi]] = \lambda w.\forall w' \in Dox(x)(w) : \phi(w')$
$w' \in Dox(x)(w)$ iff given the beliefs of x in w, w' could be the actual world

(62) **Piece 2:** Matrix K axiom:
Assertion of ϕ by a speaker S is parsed as $K_S\phi$ at LF

(63) **Piece 3:** K and exh in alternatives:

One cannot delete K and exh from alternatives.

In order to put together the pruning proposal with the grammatical theory of ignorance implicatures, we will need to introduce a modification of **Piece 2**. The reason why this modification is needed is that, as far as we can see, there is no parse of a sentence such as 'Everyone is A or B' allowed by the original version of Meyer's theory such that ignorance inferences are derived with *ALT(all-or)* alternatives, and distributive inferences are derived with *ALT(or) alternatives*: the existence of such a parse is necessary to capture the empirical pattern of ignorance and distributive inferences of such sentences under the informativeness-based pruning proposal. Therefore, we propose to modify **Piece 2** minimally as in (64).

(64) **Piece 2*:** Modified Matrix K axiom:
Assertion of ϕ by a speaker S is parsed as $K_S exh(\phi) \wedge exh(K_S \phi)$ at LF

Let us first see that this modified version of Meyer's theory captures the ignorance inferences of unembedded disjunction. According to (64), the sentence in (65a) uttered by a speaker S is parsed as (65b).

(65) a. John is A or B.
b. [K_S[*exh* [John is A or B]]] & [*exh* [K_S[John is A or B]]]

As *ALT*(John is A or B) = {John is A, John is B, John is A and B}, and as *IE*(John is A or B) = {John is A and B}, the meaning of *exh*(John is A or B) = 'John is A or B and John isn't A and B'. The meaning of K_S(*exh*(John is A or B)) is thus: S believes that John is A or B and that John isn't A and B.

On the other hand, by (63), *ALT*(K_S (John is A or B)) = {K_S (John is A), K_S (John is B), K_S (John is A and B)}. All of these alternatives are innocently excludable: the meaning of *exh*(K (John is A or B)) = 'S believes that John is A or B, and S doesn't believe that John is A, and S doesn't believe that John is B, and S doesn't believe that John is A and B'.

Putting this together, the predicted meaning of (65a), parsed as (65b) is in (66):

(66) S believes that John is A or B and that John isn't A and B, and S doesn't believe that John is A, and S doesn't believe that John is B.

(66) entails that the speaker has no belief about whether or not John is A, and no belief about whether or not John is B (ignorance inferences).

Let us now see how to derive distributive and ignorance inferences with the cases of embedded disjunction. We will first see that when no alternatives are pruned, ignorance inferences are derived. According to (64), (67a) is parsed as (67b).

(67) a. Everyone is A or B.
b. [K_S[*exh* [Everyone is A or B.]]] & [*exh* [K_S[Everyone is A or B.]]]

As *ALT*(Everyone is A or B) = {Everyone is A, Everyone is B, Everyone is A and B, Someone is A or B, Someone is A, Someone is B, Someone is A and B}, and as *IE*(Everyone is A or B) = {Everyone is A and B, Someone is A and B}, the meaning of *exh*(John is A or B) = 'Everyone is A or B and no one is A and B' (cf. Section 7.3.1).

The meaning of K_S(*exh*(Everyone is A or B)) is thus: S believes that everyone is A or B and that no one is A and B.

On the other hand, by (63), $ALT(K_S$ (Everyone is A or B)) = {K_S (Everyone is A), K_S (Everyone is B), K_S (Everyone is A and B), K_S (Someone is A or B), K_S (Someone is A), K_S (Someone is B), K_S (Someone is A and B)}. All of these alternatives apart from 'K_S (Someone is A or B)'[8] are innocently excludable: the meaning of *exh*(K (Everyone is A or B)) = '*S* believes that everyone is A or B, and *S* doesn't believe that everyone is A, and *S* doesn't believe that everyone is B, and *S* doesn't believe that someone is A, and *S* doesn't believe that someone is B, and *S* doesn't believe that someone is A and B, and *S* doesn't believe that everyone is A and B'.

Putting this together, the predicted meaning of (67a), parsed as (67b) is in (68):

(68) *S* believes that everyone is A or B and that no one is A and B, and *S* doesn't believe that everyone is A, and *S* doesn't believe that everyone is B, and *S* doesn't believe that someone is A, and *S* doesn't believe that someone is B.

(68) entails that the speaker has no belief about whether or not everyone is A, no belief about whether or not everyone is B, no belief about whether or not someone is A, no belief about whether or not someone is B (ignorance inferences).

Let us now see that when existential replacement alternatives are pruned[9], distributive inferences are derived. In order to apply the informativeness constraint on pruning, we will need to make the connecting assumption in (69).

(69) **Connecting assumption.**
Let A, B be alternatives of a sentence S uttered by a speaker x, $P(A)$, $P(B)$ the likelihoods of their propositional meaning, and $P(K_x(A))$, $P(K_x(A))$ the likelihoods that the speaker believes A (B) according to the listener; we assume that $P(A) > P(B)$ iff $P(K_x(A)) > P(K_x(B))$.

We can thus import the pruning mechanism according to which the least informative alternatives are pruned. When the sentence in (67a) is parsed as (67b), and the least informative alternatives are pruned from *ALT*(Everyone is A or B) and from $ALT(K_S$ (Everyone is A or B)) (i.e. when the existential replacement alternatives are pruned), *IE*(Everyone is A or B) = {Everyone is A, Everyone is B, Someone is A and B, Everyone is A and B}, and $IE(K_S$ (Everyone is A or B)) = { K_S (Everyone is A), K_S (Everyone is B), K_S (Someone is A and B), K_S (Everyone is A and B)}. The resulting meaning of (67a) will thus be in (70), i.e. we derive distributive inferences.

(70) *S* believes that everyone is A or B and that no one is A and B and that not everyone is A and that not everyone is B.

In this section, we have worked out one way the informativeness-based pruning proposal can be put together with a grammatical theory of ignorance implicature derivation. We do not intend to claim that this is the best way to achieve this: evaluating this is left for future work.

[8] K_S (Someone is A or B) is entailed by K_S (Everyone is A or B) when the domain of individuals is non-empty.

[9] We are not going to assume that the alternative 'Some are A and B' is pruned: whether or not it is pruned is irrelevant for our purposes.

7.11 Empirical challenges

In this section, we will discuss four empirical challenges for the idea that the deviance of sentences such as (37) and (39) is due to ignorance inferences. We address some of them, and leave others for future research.

7.11.1 Modal contrast

There is a contrast between the possibility and necessity modal that is relevant for the phenomenon in question: while (71a) sounds perfectly fine, (71b) is deviant.

(71) a. Each of these three girls might be Mary, Susan, or Jane.
b. #Each of these three girls must be Mary, Susan, or Jane.

What inferences are triggered by sentences in (71)? To answer that, let us examine the inferences of the minimally different sentences in (72), which are both well-formed:

(72) a. Each of these three girls might be from France, Germany, or Spain.
b. Each of these three girls must be from France, Germany, or Spain.

Intuitively, both of the sentences in (72) trigger free choice inferences in (73).

(73) Girl 1 might be from France, Girl 2 might be from France, Girl 3 might be from France (likewise, Germany, Spain...)

How about ignorance inferences triggered by (72a) and (72b)? In particular, do they trigger the ignorance inferences in (74)?

(74) The speaker is ignorant about whether each of these three girls is from France. (likewise, from Germany, Spain...)

It is not clear to us what the empirical facts are about the ignorance inferences in (74). If one could however demonstrate that (72b) but not (72a) triggers the ignorance inferences in (74), one could capture the contrast in (71): in that case, (71b) but not (71a) would trigger the ignorance inferences in (75), which contradict contextual knowledge.

(75) The speaker is ignorant about whether each of these three girls is Mary. (likewise, Susan, Jane...)

Establishing whether this is a viable approach for the explanation of the contrast in (71) is left for future work.

7.11.2 Larger domain examples

Consider the deviant sentence in (76).

(76) #Each of the twenty girls in this photo is Mary's daughter Lisa or one of her classmates.

If the sentence in (76) triggered ignorance inferences, we could explain its deviance in the same way as we did for the sentences (37) and (39). However, while we provided independent evidence that sentences such as (37) and (39) should trigger ignorance inferences in Section 7.2, such evidence is lacking for (76). In particular, we have seen universally quantified sentences with the cardinality of restrictor and the number of disjuncts as in (76), which clearly trigger distributive and not ignorance inferences (cf. (4)).

(4) All 20 of Mary's friends are French or Spanish.

We would like to point to an alternative approach for the deviance of (76), which is nonetheless in the same spirit as the current proposal. Spector (2018) observes that sentences such as (4) trigger not only the distributive inferences according to which at least one of the twenty girls is French, and at least one is Spanish, but also an inference about how many of the twenty girls (approximately) are French, and how many are Spanish (we will refer to this in the continuation as the *distribution estimate inference*). The content of this inference for a sentence such as (4) seems to be that there is approximately as many of the twenty girls are French as those who are Spanish.

If this distribution estimate inference is blind to common knowledge too, we can explain the deviance of (76). In other words, (76) would be deviant because it triggers the distribution estimate inferences in (77).

(77) Approximately the same number of the girls in the photo are Mary's daughter Lisa as the number of girls in the photo who are her classmates.

It is in fact extremely straightforward to extend any of the approaches to implicatures derivation discussed in this paper to capture the *distribution estimate inference*: the only necessary component is to assume that sentences such as (4) activate not only the alternatatives in which the universal quantifier is substituted with an existential, but also the alternatives in which the universal quantifier is substituted with the full range of numeric expressions between 1 and the n, with n being the cardinality of the restrictor. One could then say that, even though we prune the least informative alternatives from ALT((4)), the alternatives such as for instance those in (78) are sufficiently informative not to be pruned. They can all be negated consistently with (4), giving rise to inference in (79).

(78) At least 12 of Mary's friends are French, At least 12 of Mary's friends are Spanish,
...
At least 19 of Mary's friends are French, At least 19 of Mary's friends are Spanish,
All 20 of Mary's friends are French, All 20 of Mary's friends are Spanish

(79) At least 8 of Mary's friends are Spanish and at least 8 are French.

See also Appendix C for the demonstration that assuming that quantified sentences activate a full range of numerical replacement alternatives is perfectly compatible with the pruning proposal we put forward in the first part of the paper.

7.11.3 Downward-entailing contexts

We have so far not discussed what happens when sentences such as (37) and (39) are embedded under a downward-entailing operator. (80) is an example that illustrates this. The intuitions about (80) seem to be less stable, but at least some speakers find the sentence deviant.

(80) ?It's not the case that both of these girls are Susan or Jane.

Interestingly, the empirical pattern is similar to that observed with the cases discussed in Magri (2009a) in relation to the blind scalar implicature proposal. Consider (81), which is deviant just like (58), repeated here in (82).

(81) #It's not the case that some Italians come from a cold country.

(82) #Some Italians come from a warm country.

To explain the deviance of (81), Magri (2009a) proposes that implicatures are in the cases such as (81) derived *locally* instead of *globally*, that is to say, that the implicatures are derived at the embedded level, below negation, rather than at the matrix level.

We could likewise propose that (80) is deviant because ignorance inferences are derived locally. This is not an entirely satisfactory solution however: while we do have empirical evidence that scalar implicatures can be sometimes derived locally in the scope of downward-entailing operators such as negation, such evidence is currently lacking for ignorance inferences.

It should be noted however that it is possible to construct deviant cases with the modified numeral *at least* in downward-entailing contexts (recall that this modified numeral triggers ignorance inferences in upward-entailing contexts which might result in deviance of the sentence, cf. (57)). Consider the slightly adapted scenario from Buccola and Haida (2018) in (83).

(83) Context: Ann played a card game in which, given the rules, the final score is always an even number of points. According to the rules, if a person scores 2 or 4 points, they get a small prize, and if they score 6 or more, they get a big prize. Carl believes that Ann scored at least 6 points; Bob however sees that Ann is awarded a small prize, and reports to Carl:

a. ?Ann got a small prize, so it can't be the case that she scored at least 5 points.
b. Ann got a small prize, so it can't be the case that she scored at least 6 points.

It is thus possible that whatever is responsible for the deviance of (83a) (be it local derivation of ignorance inferences or something else) is also responsible for the deviance of (80).

7.11.4 Guessing contexts

Scalar and ignorance implicatures display different behaviors in guessing contexts (Fox, 2014). Adapting an example from Fox, 2014, consider a contest scenario in which there are 5 boxes, some of which contain money, and a contestant obtains the money if they guess correctly all of the boxes in which the money is. The host of the contest knows exactly which boxes contain money, and they are allowed to give a cue to the contestant about it, without being allowed to disclose the complete answer to the contestant. Suppose now that the host says (84).

(84) The money is in Box 1 or Box 4.

Fox (2014) argues that in this scenario (84) triggers the inference that the money is not in both Box 1 and Box 4 (the scalar implicature), but that it does not trigger the inference that the speaker is ignorant about whether the money is in Box 1 and about whether the money is in Box 4. Fox (2014) proposes that ignorance inferences of (84) are not derived because the maxim of quantity is not active in such a competition scenario (the speaker is not expected to disclose all of the relevant information in this scenario, unlike in ordinary conversation scenarios).

The lack of ignorance inferences coupled with the presence of scalar implicature in (84) raises a general challenge for the grammatical theories of ignorance inferences, which do not rely on the maxim of quantity for ignorance inferences derivation. Furthermore, the absence of ignorance inference in guessing contexts also raises a more specific challenge for the proposal according to which the deviance of (37) and (39) is due to ignorance inferences. The reason for this more specific challenge is that there doesn't seem to be an improvement of sentences like (37) and (39) even in guessing scenarios[10].

(85) Here are 3 writers: Tolstoy, Zola, and Rowling. #Each of them wrote Anna Karenina, Germinal, or Harry Potter. Please tell me which of the three writers wrote which book.

We will now attempt to address both of these challenges. We propose that the ignorance inferences of (84) are absent in a guessing scenario because the modal in the scope of which the utterance is interpreted changes. In particular, while in ordinary conversation scenarios this modal is K, with the meaning akin to 'the speaker believes', in guessing contexts this modal has the meaning akin to 'the speaker is allowed to say', which we will write as $\diamond SAY_x$, with the semantics in (86).[11]

(86) $[[\diamond SAY_x \phi]] = \lambda w.\exists w' \in CB(w) : \phi$ is entailed by what x said in w'
$w' \in CB(w)$ iff all of the regulations from w are true in w'

In other words, we propose a final modification in (87) of the Matrix K axiom from Meyer's theory.

(87) **Piece 2* final:** Final modification of the Matrix K axiom:
Assertion of ϕ by a speaker S is parsed as $K_S exh(\phi) \wedge exh(K_S\phi)$ at LF in ordinary conversation contexts, and as $\diamond SAY_x exh(\phi) \wedge exh(\diamond SAY_x \phi)$ in guessing contexts.

Let us see that this correctly predicts no ignorance inferences of (84), while still predicting the deviance of (37) and (39) in guessing contexts (cf. (85)).

The predicted meaning for (84), parsed as (88), is in (89): no ignorance inferences result from this parse, only the inferences that the speaker is not allowed to say that the money is in Box 1 and that he is not allowed to say that the money is in Box 4, while he is allowed to say that the money is in Box 1 or Box 4 and that it is not in both Box 1 and Box 4.

(88) $\diamond SAY_x exh$(The money is in Box 1 or Box 4) $\wedge$ $exh(\diamond SAY_x$The money is in Box 1 or Box 4)

(89) $\neg\diamond SAY_x$ the money is in Box 1 and $\neg\diamond SAY_x$ the money is in Box and $\neg\diamond SAY_x$ money is in Box 1 and Box 4, and $\diamond SAY_x$ that the money is in Box 1 or Box 4 and $\diamond SAY_x$ the money is not in both Box 1 and Box 4.

Let us now move to the case (37), to demonstrate that even with the guessing context parse in (90), the contradictory inferences are predicted.

[10] Thanks to an anonymous conference reviewer for this observation.

[11] Other possibilities are conceivable here: for instance, Benjamin Spector (p.c.) suggests that in certain guessing contexts the modal in the scope of which the utterance is interpreted might have the meaning more akin to 'the speaker wants to say'.

(90) $\diamond SAY_x exh$(Each of these girls is Mary, Susan, or Jane)
$\wedge\ exh(\diamond SAY_x$Each of these girls is Mary, Susan, or Jane)

Assuming as before that existential replacement alternatives are activated by a sentence such as (37), the meaning of (37) parsed as (90) is (abbreviated) in (91):

(91) $\neg\diamond SAY_x$ each of these girls is Mary $\neg\diamond SAY_x$ some of these girls is Mary.... and $\diamond SAY_x$ each of these girls is Mary, Susan, or Jane....

The problem with this meaning is that '$\neg\diamond SAY_x$ some of these girls is Mary' and '$\diamond SAY_x$ each of these girls is Mary, Susan, or Jane' contextually contradict each other (because 'each of these girls is Mary, Susan, or Jane' contextually entails 'some of these girls is Mary'). (37) and (39) are thus predicted to be unacceptable even in guessing contexts if the modification of the Matrix K axiom proposed in (87) is on the right track.

7.12 Concluding remarks

In this paper, two novel empirical puzzles have been explored.

The first puzzle taught us that quantified sentences with embedded disjunction trigger ignorance inferences which are sensitive in some way to the informativeness of the utterance (as evidenced by the effect of the domain size and the number of disjuncts on whether the ignorance or the distributive inferences are derived). The account we have put forward to capture this effect is that pruning of alternatives is sensitive to how much information the alternative carries over and above the original utterance: the more informative the alternative is, the more likely it is to be kept in the alternative set in the computation of implicatures. Importantly, even if the specifics of the pruning account turn out to be incorrect, the data pattern that the account aims to capture strongly suggests that informativeness other than logical or contextual entailment plays a role in some way in implicature computation. Strikingly, the computation of informativeness which plays a role in implicature derivation seems to be blind to (at least) some aspects of contextual knowledge (i.e. there has to be some level of modularity when informativeness is calculated). We consider these two conclusions — that probabilistic informativeness might be at heart of implicature derivation, and that it is computed in a modular way — to be the most important contribution of the first part of the paper.

The second puzzle explores a novel case of deviance, which we have argued to be caused by ignorance inferences. Importantly, if the proposed account is on the right track, ignorance inferences need to be derived blindly to contextual knowledge (much like scalar implicatures have been argued to be derived blindly to contextual knowledge by Magri, 2009b). If ignorance inferences are derived blindly to contextual knowledge, we have an empirical argument to model them as semantic inferences on a par with scalar implicatures in exhaustification approaches, rather than as pragmatic inferences (cf. Meyer, 2013, 2014, and Section 7.10).

To say that a component of the mechanism for ignorance inferences derivation which calculates informativeness of alternative utterances is blind to contextual knowledge does not entail that contextual knowledge plays no role whatsoever in the derivation of ignorance inferences. There is in fact empirical evidence that contextual knowledge plays some role in it (lack of ignorance inferences in guessing contexts, cf. Fox, 2014). We have proposed that in this particular case, the context influences the covert modal in the scope of which the utterance is interpreted. It is an open question whether context might influence ignorance inference derivation is some additional ways, and how this influence is to be modelled.

Zooming into the features of the pruning proposal, we have proposed to distinguish between relevance-based pruning and informativeness-based pruning, assuming that they play partly complementing roles in restricting the formal alternative set for implicature derivation: the set of formal alternatives that is used in the implicature derivation is the set of formal alternatives which are both contextually relevant and sufficiently informative. Importantly, relevance-based pruning is essentially connected to contextual knowledge, while we have argued that the informativeness-based pruning is blind to at least some aspects of it. This fact may suggest that the two pruning procedures are of very different nature. One possibility is that there are in fact two routes to implicatures: one truly pragmatic (and thus domain-general), and the other grammatical (modular). In other words, some instances of people's implicature derivation might be truly the result of Gricean-like reasoning, and others the result of a grammar-internal (modular) calculation. Relevance-based pruning might thus be a component of the first route, and informativeness-based pruning a component of the second route.

A straightforward objection to the two routes idea is why the cases of deviance due to blind implicatures ever occur. A possible reason why is that the grammatical calculation of implicatures might be less costly than the pragmatic one (in terms of time or cognitive ressources need to complete the derivation), and is thus a default one: there might be certain contextual requirements to be met for someone to engage in the pragmatic calculation of implicatures.

Clearly, the two routes idea is very speculative, and requires much further research.

Appendix A: Probabilities

Fact 1

We show that for any $p_A \in]0,1[$, for any $m,n \in \mathbb{N}$, if $m < n$, the following is true:
$1-(1-p_A)^m < 1-(1-p_A)^n$.

When $p_A \in]0,1[$, the function $f : n \mapsto p_A^n$ is a decreasing function: the bigger n, the smaller p_A^n.
Note that the smaller p_A^n, the bigger $1-p_A^n$.

Thus for any $m,n \in \mathbb{N}$, if $m < n$, it holds that $1-(1-p_A)^m < 1-(1-p_A)^n$.

Fact 2

We show that for any $p_A \in]0,1[$, for any $n \in \mathbb{N}$ such that $n > 1$, the following is true:
$1-(1-p_A)^n > p_A^n$.

Let $p_A \in]0,1[$, and $n > 1$. Then: $p_A > p_A^n$ and $(1-p_A) > (1-p_A)^n$. Summing up then: $1 = p_A + (1-p_A) > p_A^n + (1-p_A)^n$, and so $1-(1-p_A)^n > p_A^n$.

Fact 3

Suppose A_1, A_2, ..., A_k are predicates such that for any relevant individual, the probability that they satisfy A_1 equals the probability that they satisfy A_2, ..., equals $p \in]0,1[$. Assume further that whether or not an individual satisfies some predicate is independent of whether this individual satisfies other predicates, and of which predicates other individuals satisfy. For a domain size $n \in \mathbb{N}$ then, *P(Some of the n individuals is A_1 |All n individuals are A_1 or A_2...or A_k)* = $P(\exists i, i \in A_1 \mid \forall i, i \in A_1 \cup A_2 ... \cup A_k)$ is equivalent to $1-(1-\frac{p}{1-(1-p)^k})^n$. Let us see why.

For some individual x_j, the probability of $x_j \in A_1 \cup A_2 ... \cup A_k$ is:
(i) $P(x_j \in A_1 \cup A_2 ... \cup A_k) = 1 - (1-p)^k$

Let us now calculate the probability of $x_j \in A_1$ conditioned on knowing that in the domain of n individuals, $\forall x \in A_1 \cup A_2 ... \cup A_k$.

$P(x_j \in A_1 \mid \forall x, x \in A_1 \cup A_2 ... \cup A_k) =$
$P(x_j \in A_1 \mid x_1 \in A_1 \cup A_2 ... \cup A_k \wedge x_2 \in A_1 \cup A_2 ... \cup A_k ... \wedge x_n \in A_1 \cup A_2 ... \cup A_k) =$
(by independence assumption)
$P(x_j \in A_1 \mid x_j \in A_1 \cup A_2 ... \cup A_k) =$
(by Bayes rule)
$\frac{P(x_j \in A_1 \wedge x_j \in A_1 \cup A_2 ... \cup A_k)}{P(x_j \in A_1 \cup A_2 ... \cup A_k)} =$
(because $A_1 \subset A_1 \cup A_2 ... \cup A_k$)
$\frac{P(x_j \in A_1)}{P(x_j \in A_1 \cup A_2 ... \cup A_k)} =$
(due to (i))
$\frac{p}{1-(1-p)^k}$

Given this, the probability of $x_j \notin A_1$ conditioned on knowing that in the domain of n individuals, $\forall x \in A_1 \cup A_2 ... \cup A_k$ is:
$P(x_j \notin A_1 \mid \forall x, x \in A_1 \cup A_2 ... \cup A_k) = 1 - \frac{p}{1-(1-p)^k}$.

Given the assumption that each of the n individuals has the same probabilities $p_{A_1}, p_{A_2}, ... p_{A_k}$, of being in $A_1, A_2, ... A_k$, it follows that: $P(\forall x, x \notin A_1 \mid \forall x, x \in A_1 \cup A_2 ... \cup A_k) = (1 - \frac{p}{1-(1-p)^k})^n$

Given that:
$P(\exists x, x \in A_1 \mid \forall x, x \in A_1 \cup A_2 ... \cup A_k) =$
$1 - P(\forall x, x \notin A_1 \mid \forall x, x \in A_1 \cup A_2 ... \cup A_k)$, it follows that:
$P(\exists x, x \in A_1 \mid \forall x, x \in A_1 \cup A_2 ... \cup A_k) = 1 - (1 - \frac{p}{1-(1-p)^k})^n$.

We now show that for any $p \in]0,1[$, for any $h, k, n \in \mathbb{N}^*$, if $h < k$, the following holds: $1 - (1 - \frac{p}{1-(1-p)^h})^n > 1 - (1 - \frac{p}{1-(1-p)^k})^n$.
As $h < k$, $1-(1-p)^h < 1-(1-p)^k$.
Thus $\frac{p}{1-(1-p)^h} > \frac{p}{1-(1-p)^k}$.
Thus $1 - \frac{p}{1-(1-p)^h} < 1 - \frac{p}{1-(1-p)^k}$.
Thus for any $n \in \mathbb{N}^*$ it holds that $(1 - \frac{p}{1-(1-p)^h})^n < (1 - \frac{p}{1-(1-p)^k})^n$.
This further means that:
$1 - (1 - \frac{p}{1-(1-p)^h})^n > 1 - (1 - \frac{p}{1-(1-p)^k})^n$

Hence for any for any $p \in]0,1[$, for any $h, k, n \in \mathbb{N}^*$, if $h < k$, the following holds:
$1 - (1 - \frac{p}{1-(1-p)^h})^n > 1 - (1 - \frac{p}{1-(1-p)^k})^n$.

Appendix B: Recursive exhaustification and distributive inferences

Crnič et al. (2015)

There is an alternative way to derive distributive inferences in the exhaustification-based framework **if one assumes that the exhaustifying operator can apply recursively**, argued for in Crnič et al. (2015). They provide experimental results showing that sentences such as (92a) trigger the inference in (92b) without necessarily triggering the inference in (92c). This suggests that negating the alternatives 'Every box contains an A' and 'Every box contains a B' (cf. Section 7.3.1), is not the way (or at least not the only way) people derive distributive inferences.

(92) a. Every box contains an A or a B.
b. Some box contains an A and some box contains a B.
c. Not every box contains an A and not every box contains a B.

In order to derive the inferences in (92b) without deriving the inferences in (92c), Crnič et al. (2015) propose that the exhaustification operator applies twice in a sentence such as (92a). More specifically, they propose that the logical form of (92a) is (93).

(93) *exh* [Every box$_x$ *exh* [x contains an A or a B]]

According to this parse, the alternatives on which the matrix *exh* operates are in (94):

(94) a. [Every box$_x$ *exh* [x contains an A]] = Every box contains an A and not a B
b. [Every box$_x$ *exh* [x contains an B]] = Every box contains a B and not an A
c. [Every box$_x$ *exh* [x contains an A and a B]] = Every box contains an A and a B

All of these alternatives can be negated consistently with (95a), i.e. the original proposition which is an argument to the matrix *exh*, whose meaning is paraphrased in (95b).

(95) a. [Every box$_x$ *exh* [x contains an A or a B]]
b. Every box contains an A or a B and not both A and B.

Negating (94a) obtains the inference in (96a), and negating (94b) the inference in (96b).

(96) a. It's not the case that every box contains an A and not a B.
⇝ Some box contains a B.
b. It's not the case that every box contains a B and not an A.
⇝ Some box contains an A.

Crucially, then, under the parse in (93), distributive inferences in (92b) are derived without the inferences in (92c).

Let us see what inferences are predicted by recursive exhaustification once we count in the inferences obtained by replacing the universal quantifier with an existential.

The alternatives activated by (92a), parsed with two exhaustification operators as in (93), are in (97):

(97) a. [Every box$_x$ *exh* [x contains an A]] = Every box contains an A and not a B
b. [Every box$_x$ *exh* [x contains an B]] = Every box contains a B and not an A
c. [Every box$_x$ *exh* [x contains an A and a B]] = Every box contains an A and a B
d. [Some box$_x$ *exh* [x contains an A]] = Some box contains an A and not a B
e. [Some box$_x$ *exh* [x contains an B]] = Some box contains a B and not an A
f. [Some box$_x$ *exh* [x contains an A or a B]] = Some box contains an A or a B and not both A and B
g. [Some box$_x$ *exh* [x contains an A and a B]] = Some box contains an A and a B

Now, which are the maximal sets of alternatives from (97) which can be negated consistently with (95a), the original proposition which is an argument to the matrix *exh*, whose meaning is paraphrased in (95b)?

Two following sets of alternatives are the maximal sets which can be negated consistently with (95a):

(98) a. Every box contains an A and not a B, Every box contains a B and not an A, Some box contains an A and a B, Every box contains an A and a B

b. Every box contains an A and not a B, Some box contains an A and not a B, Some box contains an A and a B, Every box contains an A and a B

c. Every box contains a B and not an A, Some box contains a B and not an A, Some box contains an A and a B, Every box contains an A and a B

In other words, the only innocently excludable alternatives are (97g) and (97c). Under the same assumption as before about ignorance inferences (i.e. that they are derived about all of the alternatives which are neither innocently excludable nor entailed by the utterance), we derive the ignorance inferences about the alternatives in (97a), (97b), (97d) and (97e). The alternative in (97f) is entailed by (95a).

Crucially, then, we have the exact same situation with recursive approach to distributive inferences as it was the case with standard approach: admitting the alternatives obtained by the replacement of the universal quantifier with an existential derives ignorance inferences and doesn't derive distributive inferences; not admitting them derives distributive inferences and not ignorance inferences.

This means that the proposal in (33) can capture the cardinality of the restrictor effect even under the recursive exhaustification approach: intuitively, the reason is that the larger the domain size, the less informative (99a) given (95b) ((99a) given (95b) is less informative than (99b) given (95b) whenever $n > 1$). In other words, the larger the domain size, the more likely we will be to prune the alternatives such as (99a) and end up with the set of alternatives as in Table 7.8 rather than as in Table 7.7, deriving distributive rather than ignorance inferences.

(99) a. Some of the n boxes contains an A and not a B

b. All of the n boxes contain an A and not a B

Alternatives of (95a)	Inferences
[Every box$_x$ *exh* [x contains an A]]	**ignorance**
[Every box$_x$ *exh* [x contains an B]]	**ignorance**
[Every box$_x$ *exh* [x contains an A and a B]]	Not every box contains an A and a B
[Some box$_x$ *exh* [x contains an A]]	**ignorance**
[Some box$_x$ *exh* [x contains an B]]	**ignorance**
[Some box$_x$ *exh* [x contains an A or a B]]	(entailed by (95a))
[Some box$_x$ *exh* [x contains an A and a B]]	No box contains an A and a B

Table 7.7: Inferences of (5) — no alternatives pruned

Alternatives of (95a)	Inferences
[Every box$_x$ *exh* [x contains an A]]	Some box contains a B
[Every box$_x$ *exh* [x contains an B]]	Some box contains an A
[Every box$_x$ *exh* [x contains an A and a B]]	Not every box contains an A and a B
~~[Some box$_x$ *exh* [x contains an A]]~~	—
~~[Some box$_x$ *exh* [x contains an B]]~~	—
~~[Some box$_x$ *exh* [x contains an A or a B]]~~	(entailed by (95a))
[Some box$_x$ *exh* [x contains an A and a B]]	No box contains an A and a B

Table 7.8: Inferences of (4) — the least informative alternatives pruned

Bar-Lev and Fox (2016)

There is yet another way to derive distributive inferences with recursive exhaustification, proposed by Bar-Lev and Fox (2016), as reported in Gotzner and Romoli (2017). It consists in applying recursively the exhaustifying operator at the matrix position, assuming that sentences with the disjunction embedded in the scope of a universal quantifier activate the alternatives as in *ALT(all-or)*, i.e. the set of alternatives which include the existential replacement alternatives.

We have already established in Section 7.3.1 that if a sentence such as (100) activates the alternatives as in *ALT(all-or)*, if no alternatives are pruned as (100) is parsed as (101), ignorance inferences are derived and not distributive inferences.

(100) Everyone is A or B.

(101) [*exh* [Everyone is A or B]]

However, as Bar-Lev and Fox (2016) demonstrate, if the sentence in (100) is parsed as (102), distributive inferences are derived.

(102) [*exh* [*exh* [Everyone is A or B]]]

Let us see why. The alternatives of the topmost *exh* are thus[12]:

(103) a. *exh* [Everyone is A] = Everyone is A and no one is B
b. *exh* [Everyone is B] = Everyone is B and no one is A
c. *exh* [Someone is A] = Someone is A and not everyone is A and no one is B
d. *exh* [Someone is B] = Someone is B and not everyone is B and no one is A

All of these alternatives are innocently excludable: negating them derives the distributive inferences in (104).

(104) Someone is A and someone is B.

In other words, the situation is the following: considering that sentences with the disjunction embedded in the scope of a universal quantifier activate the alternatives as in *ALT(all-or)*, assuming that recursive exhaustification at the matrix level is possible, exhaustifying a sentence

[12]There are also the alternatives *exh*(Someone is A or B), *exh*(Someone is A and B), *exh*(Everyone is A and B); we are ignoring them for simplicity because they don't play a role in the derivation of distributive and ignorance inferences.

like (100) once derives ignorance inferences via the basic maxim of quantity, and exhaustifying it twice derives distributive inferences.[13]

If this is indeed the way distributive and ignorance inferences for sentences like (100) are derived in language, one can put forward a proposal alternative to pruning to capture the empirical pattern of distributive and ignorance inferences, one that would possibly relate the informativeness of a sentence with a propensity to parse it with recursive matrix exhaustification. The idea in brief would be that, given that (105a) is more informative than (105b), and that (106a) is more informative than (106b), we are more likely to parse (105a) with double matrix exhaustification as compared to (105b), and likewise (106a) as compared to (106b).[14] In other words, the disambiguation process (between a parse with a single *exh* and parse with double *exh*) would have to be guided by the informativeness of a sentence. In order to capture the full set of data presented in this paper, however, one would need to propose that the calculation of informativeness which guides the disambiguation is not sensitive to contextual knowledge.

(105) a. All of the 20 individuals are A or B.
b. Both individuals are A or B.

(106) a. All four individuals are A or B.
b. All four individuals are A, B, C, or D.

Appendix C

We have assumed that sentences such as (107a) activate the alternatives obtained by dropping one of the disjuncts combined with the replacement of the universal quantifier with an existential quantifier *some* (=*at least one*). However, one might argue that the alternative set is in fact even larger, and includes not only the replacement of the universal quantifier by *some*, but by an entire range of numerical expressions, as in (107b).

(107) a. All of the n individuals are A or B.
b. Some of the n individuals are A, Some of the n individuals are B, At least 2 of the n individuals are A, At least 2 of the n individuals are B,..., At least $n-1$ of the n individuals are A, At least $n-1$ of the n individuals are B, All of the n individuals are A, All of the n individuals are B

The pruning account works even with this set of alternatives: let us see why. What we need to show is (i) that when no alternatives are pruned from (107b), ignorance inferences are derived for the sentence (107a); and (ii) that when the alternative in (108) are pruned from (107b), distributive inferences are derived for (107a).

(108) Some (=at least 1) of the n people is A, Some (=at least 1) of the n people is B

Let us first show (i): that when no alternatives are pruned from (107b), ignorance inferences are derived for the sentence (107a).
Note that for a sentence (107a), there are $2n$ alternatives in (107b) (i.e. there are $2n$ of the al-

[13]If one adopts the grammatical approach to implicatures from Meyer, 2013, if the sentence (100) activates the alternatives $ALT(all-or)$, ignorance inferences are derived by exhaustifying once above the K operator, and distributive inferences are derived by exhaustifying twice below the K operator.

[14]This is true under the same assumptions as elsewhere in the paper.

ternatives obtained by dropping one of the disjuncts (with or without the replacement of the universal quantifier by other numerical expressions).

For $n > 0$, the sentence (107a) entails that there is $x > 0$ such that the following holds: at least x individuals are A and at least $n - x$ individuals are B.

Taking into account entailment relations among alternatives, this means that (107a) entails that at least n alternatives from (107b) are true.

This further means that at most n alternatives (107b) can be negated consistently with (107a).

This further means that, if there is a set of n alternatives which can be negated consistently with (107a) such that it doesn't contain an alternative in (109a), the alternative in (109a) is not innocently excludable (and thus ignorance inferences will be derived about it). Such a set exists: it's in (109b). Similarly, the alternative in (110a) is not innocently excludable either (because there are n alternatives in (110b) can be negated consistently with (107a), and (110a) is not one of them). This basically means that, if (107a) activates the full range of numerical alternatives (and none of them are pruned), alternatives in (109a) and (110a) will not be innocently excludable, and thus distributive inferences will not be derived, but rather ignorance inferences about (109a) and (110a).

(109) a. All of the n individuals are A.
b. Some of the n individuals are B, At least 2 of the n individuals are B,..., At least $n-1$ of the n individuals are B, All of the n individuals are B

(110) a. All of the n individuals are B.
b. Some of the n individuals are A, At least 2 of the n individuals are A,..., At least $n-1$ of the n individuals are A, All of the n individuals are A

Let us now show (ii): that when the alternative in (108) are pruned from (107b), distributive inferences are derived for (107a).

As before, the number of alternatives from (107b) which can be negated consistently with (107a) is at most n.

It is possible to come up with a set of alternatives of n members which can be negated consistently with the (107a) and which contain both (109a) and (110a), for instance, (111) (negating the alternatives in (111) will give rise to the inference 'At least x individuals are B and at least $n-x$ individuals are A').

(111) At least $x+1$ individuals are B, At least $x+2$ individuals are B,..., At least $n-1$ individuals are B, All of the n individuals are B, At least $(n-x)+1$ individuals are A, At least $(n-x)+2$ individuals are A,..., At least $n-1$ individuals are A, All of the n individuals are A

However, because the alternatives in (108) are pruned, it is no longer possible to come up with a set of alternatives of n members which can be negated consistently with (107a) which do not contain (109a) and (110a).

This means that the alternatives (109a) and (110a) are innocently excludable when alternatives in (108) are pruned from the alternative set in (107b), and that distributive inferences are derived for a sentence (107a).

Because of (i) and (ii), the proposal from (33) applies without any further assumptions even if we assume that sentences such as (107a) activate all of the alternatives from (107b).

Acknowledgements

Many thanks to Emmanuel Chemla and Benjamin Spector for important input throughout the development of this work, as well as to Amir Anvari, Moshe E. Bar-Lev, Marc Benhamou, Brian Buccola, Gennaro Chierchia, Rachel Dudley, Émile Enguehard, Diego Fienmann, Danny Fox, Manuel Križ, Mora Maldonado, Jacopo Romoli, Philippe Schlenker, and Andrés Soria Ruiz for very helpful comments and discussions. I am also grateful to audiences of SALT28, SuB23, and ETAPS workshop for their comments and suggestions.

Chapter 8

Quantifier spreading in child language

This chapter is a reproduction of the following article:

Denić, Milica and Emmanuel Chemla. 2019 (to appear). Quantifier spreading in child language as distributive inferences. *Linguistic Inquiry*

Abstract

We focus on a child language phenomenon called quantifier spreading, whereby children judge sentences such as 'Every girl took an apple' to be false in situations where there is an apple not taken (Inhelder and Piaget, 1964, a.o.). These are typically interpreted as "mistakes". We show however that formal semantic theories actually make the prediction that this is an accurate response, and that it is the adults' responses that need to be explained. This has consequences about these theories (ie. whether they are good theories of child and/or adult language competence) and this approach opens the way for a systematic analysis of child language, which is not always studied with as detailed models as adults.

8.1 Missing implicatures in adult language

8.1.1 Disjunctions: free choice and distributive inferences

Disjunctions in the scope of a possibility modal trigger so-called *free choice inferences*: (1a) gives rise to the inference (1b) (e.g., Kamp 1973).

(1) a. John can read Article 1, Article 2, or Article 3.
    b. John can read Article 1, and he can read Article 2, and he can read Article 3.

In the literature, such inferences are typically (but not always) derived as implicatures, crucially relying on the assumption that a sentence with a disjunction activates domain alternatives. To understand what domain alternatives are, note that a disjunction can be described by the set of elements (objects or propositions) that it covers. For the disjunction in (1a) that set of elements, that we will refer to as the domain D of disjunction, would be $D = \{\text{Article 1}, \text{Article 2}, \text{Article 3}\}$. Domain alternatives of a disjunction are other disjunctions which differ from the original in that they are constructed on smaller domains $D' \subseteq D$. In other words, a sentence with a disjunction, schematically P(A1 or A2 or A3), with P(__) standing for the environment in which the disjunction is embedded, activates domain alternatives of the form: P(A1 or A2), P(A1 or A3), P(A2 or A3), P(A1), and so on. These alternatives have been argued to serve as input for a mechanism which derives free choice inferences in (1b). All details and motivation can be found in Fox (2007).

Leaving aside the technical details, assume that one can thus consider free choice inferences as evidence that sentences with disjunction activate domain alternatives. Given most theories of implicatures, say that of Chierchia et al. (2012) for concreteness, this not only explains the free choice inference in (1), it also makes predictions about the interpretation of a disjunction in the scope of a universal quantifier: a sentence such as (2a) intuitively gives rise to the inference (2b), and this is indeed a prediction of this approach; let us see why. Schematically, in (2a) the disjunction appears in the environment P(__) = *Every girl took* __, and therefore one of the domain alternatives that (2a) triggers is P(A2, A3, or A4). This alternative is stronger than the original sentence and, according to standard assumptions about implicature derivation, it therefore ends up being negated, that is, the original sentence gives rise to the inference not-P(A2, A3, or A4). One thus obtains the inferences that 'it is not the case for every girl that she took Apple 2 or Apple 3 or Apple 4' which, together with the original utterance, entails that 'some girl took Apple 1'. Similarly for every apple x in the domain, one obtains the inference that 'some girl took Apple x'. All of these together amount to the inference stated in (2b).

(2) a. Every girl took Apple 1, Apple 2, Apple 3, or Apple 4.
  b. Each of the four apples was taken by some girl.

In sum, given the alternatives evidenced by free choice inferences (cf. (1)), one may derive implicatures and judge (2a) as not true in the situation depicted in Figure 8.1. This prediction is borne out, these implicatures are called distributive inferences (cf. Spector 2006 as well as quantitative data in Crnič et al. 2015 and the experiment reported in Appendix 8.5.2).

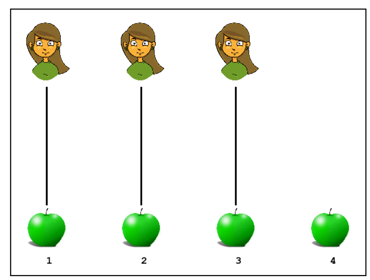


Figure 8.1: Every girl took an apple, but not every apple was taken.

8.1.2 Indefinites: free choice but no distributive inferences

Indefinite noun phrases also trigger free choice effects in the scope of a possibility modal: in a context with three salient articles, a possible reading of (3a) is (3b).[1]

(3) a. John can read an article.
  b. John can read Article 1, and he can read Article 2, and he can read Article 3.

Indefinites introduce existential quantification over a contextually supplied domain, assume for (3a) that this domain is D = {Article 1, Article 2, Article 3}. One can then postulate that the indefinite activates domain alternatives, obtained by exchanging the contextually supplied D with its subsets. For the same reasons as above, this would generate free choice inferences with indefinites. As it was done above for disjunctions, one may thus reason inductively and assume that free choice inferences observed in (3) can be taken as evidence that indefinites (and not only disjunctions) activate domain alternatives. This automatically makes a prediction for cases such as (4a), in which the indefinite appears in the scope of the universal quantifier, namely, that because of the implicatures derived due to domain alternatives of the indefinite,

[1] The indefinite in (3) can also get a specific interpretation; (3) thus has a reading in which there is a particular article which John can read, but this is not the reading that concerns us here.

(4a) could be interpreted as (4b), and therefore be judged not true in the situation depicted in Figure 8.1. Interestingly, there is no experimental or introspective evidence for such a reading in adults as shown by the experiment reported in Appendix 8.5.1.

(4) a. Every girl took an apple.
    b. Every (relevant) apple was taken by some girl.

8.1.3 Summing up: a dissociation in adult data

If free choice effects are a signal of implicatures obtained by the negation of domain alternatives, observed for both disjunctions and indefinites, then distributive inferences are predicted for both disjunctions and indefinites in the scope of a universal quantifier. However, distributive inferences are only attested for disjunctions. There are several ways to go from here. First, one may abandon unifying (implicature) theories of indefinites/disjunctions for free choice/distributive inferences. For instance, there are theories of free choice inferences that treat them as entailments rather than implicatures (cf. Aloni 2007; Barker 2010) and don't make predictions for how indefinites and disjunctions will be interpreted in the scope of a universal quantifier. Alternatively, as sketched later on in Section 8.3, one may try to amend implicature theories of free choice so that one can explain the observed dissociation (the absence of distributive inferences with indefinites). We will now discuss a child language phenomenon that may be taken as an argument for the latter way.

8.2 Children: q-spreading as distributive inferences

Strikingly, children seem to have the exact interpretation of the indefinite in the scope of a universal quantifier that is predicted, yet unattested, in adult language. That is, they would report that (4a) is false in a situation as in Figure 8.1. This phenomenon in child language has been called quantifier spreading (hereafter q-spreading), Type A error, or exhaustive pairing error. It has been extensively studied at least since Inhelder and Piaget (1964).

8.2.1 Q-spreading as distributive inferences

Given the previous discussion, an account of q-spreading is readily available: whatever is responsible for the different behaviors of indefinites and disjunctions in the scope of a universal quantifier in adult language is not operative in child language. Q-spreading in child language may thus be the result of implicatures, derived by negating domain alternatives that indefinite noun phrases ought to activate, since they give rise to free choice inferences.

Such an account is coherent with and similar in spirit to the account of Singh et al. (2016) for the conjunctive interpretation of disjunctions by children in English (cf. Tieu et al. 2017 for French, German and Japanese). The two approaches offer accounts of phenomena that are usually considered to be language performance errors in child language in terms of results of a sophisticated and legitimate inference mechanism. They thus call for a research program investigating how much of children's language performance errors are in fact mis-classified linguistic inferences.[2]

[2] Mascarenhas (2014) discusses a related research program of sorting adults' reasoning 'errors' between actual performance errors and legitimate linguistic inferences.

8.2.2 More empirical data points

We cannot review all existing accounts of q-spreading; in short, the issue has been attributed to numerous domains such as the syntax or syntax-semantics interface (e.g., Roeper and de Villiers 1991, Philip 1995, Drozd and van Loosbroek 1998, Geurts 2003), performance (Crain et al. 1996) or problems with cognitive resources and shallow processing (Freeman et al. 1982; Brooks and Sekerina 2006). Most of these however agree in that they locate the problem in the universal quantifier (its syntax, semantics, or complexity). The account we describe (hereafter implicature account) does not and is thus a distinct option which ought to be evaluated in depth. The implicature account of q-spreading also has another simple merit: a full account of children's grammar is obtained, simply as the typical account already motivated for adults, and the explanatory load is in fact moved away from the question of why children have q-spreading to why adults don't. Let us thus review how this account connects with empirical knowledge on the topic more broadly.

8.2.2.1 Developmental path

Aravind et al. (2017) conducted a longitudinal study in which children were tested four times between ages 4-7. They found that, at first, children showed little q-spreading, but that the amount of q-spreading errors increased with each subsequent testing. The same effect was found in two cross-sectional studies reported in Kang (2001) (although not in the experiment reported in Philip and Takahashi 1991). This U-shape trajectory is difficult to accommodate under accounts which blame limited cognitive resources and processing for q-spreading (for instance, Brooks and Sekerina 2006). On the other hand, it has been shown that young children start out with low rates of implicatures, (Noveck 2001, a.o), and if one could explain how they move from these low rates to high rates of implicatures, there is a potential to explain the initial increase of q-spreading errors, which are nothing else than (unwarranted) implicatures in this view.

An explanation of the increase of q-spreading interpretations over development as an increase of implicature derivation is thus tempting, but at this stage it requires some caution too. One complexity is that the typical finding that children do not show high rates of implicatures concerns implicatures requiring *scalar* alternatives. Free choice inferences and distributive inferences, if they are implicatures, are based on *domain* alternatives instead, and such inferences are more frequently derived by children (cf. Tieu et al. 2015, Singh et al. 2016, Tieu et al. 2017, and Pagliarini et al. 2018). So, the rate of derivation of these inferences is initially higher than that of regular scalar implicatures for young children, but it is not known whether they are at *ceiling* early on. If they are not at ceiling they could also increase over time, which is what would be needed to explain the developmental path for q-spreading in the current framework. Overall, a direct comparison between the rate of derivation of free choice inferences and the rate of q-spreading interpretations in different populations would be a most direct test of the implicature approach.

8.2.2.2 Contextual effects

It has been noticed that certain manipulations of experimental context can significantly reduce the amount of q-spreading (Crain et al. 1996; Philip 2011). For instance, Philip (2011), and similarly Hollebrandse (2004), report that what is considered to be a topic in the experimental context can have an influence on q-spreading. When the topic surrounding the experimental question is about the subject noun which restricts the universal quantifier, there is less q-spreading than when the topic is about the indefinite noun. An implicature account of q-spreading has the potential to capture this context dependency, since context naturally influences the derivation of implicatures. More specifically, what might be behind the reduction of quantifier spreading when the experimental topic is about the subject noun is the following: the more salient the alternative trigger (in this case the indefinite), the more likely a child (or adult) might be to generate its alternatives and proceed to the derivation of implicatures. The manipulation in Philip's and Hollebrandse's experiments might be drawing children's attention to the subject noun and away from the indefinite noun: in other words, their manipulations might be reducing the salience of the alternative trigger, and thereby reducing the impact of the alternatives.

Appendix 8.5.3 reveals one way in which context dependency can be observed in the adult counterpart of q-spreading, that is in distributive inferences with disjunction. This gives further credit to the idea that q-spreading and distributive inferences have much in common.

8.2.2.3 Superficially similar errors

Similarly, one may wonder whether q-spreading, also called type A error, is similar to other types of 'errors' children make with universally quantified sentences. Children make other errors with sentences such as (4a), and the implicature account is only able to explain q-spreading. But the other relevant errors show differences with q-spreading which may justify a separate account. First, children sometimes accept (4a) when not every girl took an apple, but all of the apples in the image were taken. But these errors (type B) have been shown to have different developmental trajectories than q-spreading. In the aforementioned study by Aravind et al. (2017), with each subsequent testing children made fewer type B errors, but showed more q-spreading. Second, children may reject (4a) when every girl took an apple, but the image also displays, for instance, a boy who took a banana. This type of error, called type C error or the perfectionist response, seems to be the least common of the three types of errors and restricted to very young children Geurts (2003), which suggests that q-spreading is of a different nature than type C errors as well.

8.2.2.4 The role of the indefinite

The current account locates the issue in the presence of the indefinite, and therefore does not make predictions in the absence of an indefinite (or an alternative trigger). However, to date the literature has focused almost exclusively on how children interpret sentences with indefinite noun phrases in the scope of a universal quantifier. If q-spreading was consistently found without indefinites, the implicature account would have to justify that children could believe that some elements, like indefinites, also trigger disjunction-like alternatives.[3]

[3] Inhelder and Piaget (1964) reported that children may reject sentences like *All the circles are blue* in a situation in which all the circles are blue but there are also blue squares. Philip and Takahashi (1991) and Takahashi (1991) showed some degree of q-spreading (i) with indefinites, (ii) with transitive sentences with null objects, as in *Every boy is driving*, with children saying false when some car has no driver, and (iii) with intransitive sentences like *Every dog is sleeping*, with children saying false when there are some extra beds in which no dog is sleeping. While one may argue that there is a null indefinite hidden in these cases, what would be more relevant is to understand, by means of systematic investigations, whether a unified account of these errors with quantifier spreading with indefinites is justified to begin with: (cf. section 2.2.3 for other types of errors children make with universally quantified sentences, but arguably then with different properties).

Conversely, the implicature account makes systematic predictions when an indefinite is present: children should have q-spreading effects with indefinites whenever adults have q-spreading-like effects with disjunctions. This is unlike existing accounts of q-spreading which may locate the problem in the universal quantifier and not in the indefinite. One could thus systematically compare the interpretation of an indefinite for children and of disjunctions for adults in new environments, such as the scope of modals and quantifiers other than universal. In particular, there are data suggesting that q-spreading does not occur with definite subjects (Drozd, 2001), and one can find at least indirect reports of tests of sentences headed by other determiners such as *both* or cardinal determiners (Takahashi, 1991; Roeper et al., 2011). At this point, it is fair to say that more systematic tests would be informative, with indefinites in various environments and with the new view that q-spreading effects may take the specific form discussed here, that of distributive inferences (and not, for example, confusion about which noun is meant to restrict the universal quantifier).

8.3 Challenges and directions concerning adults and children

Under the current view, the children situation becomes simple, as their q-spreading behavior is captured by what semanticists have proposed for adults. The adults situation is more puzzling: the existence of free choice effects and the absence of q-spreading with indefinites creates a tension. Exploring possible ways to resolve this tension is important from the child language perspective, because understanding why distributive inferences with indefinites are absent in adults is crucial for understanding what it is that children have to learn in order to achieve adult-like competence in terms of inferences generated with indefinites.

One possibility is that the current account of q-spreading is correct for children, while the analysis of free choice effects as implicatures is incorrect for adults, at least for indefinites. The theory of indefinites then, independently of that of disjunctions, would then have to provide an analysis of the free choice effects for indefinites, one that does not lead to the expectation of distributive inferences. One could then argue that children are initially mistaken and believe that indefinites activate domain alternatives, possibly due to exposure to cases such as (3) — in other words, for the same reason that some semanticists are mistaken about their theory of alternatives activated by indefinites. One remaining question would then be how and what (presumably rare) input helps them recover from this assumption and reach adult-like knowledge.

Another possibility is that the analysis of free choice effects as implicatures is entirely correct, but that distributive inferences with indefinites are blocked or masked for adults by some independent mechanism, which children are not yet aware of. Let us discuss two such possible mechanisms.

8.3.1 Q-spreading masked by domain restriction of indefinites

A salient difference between indefinites and disjunctions is that, for indefinites, the relevant domain of individuals, also known as domain restriction for quantified phrases more generally (cf. von Fintel 1994, a.o.) is left implicit, and thus potentially subject to more variability. This could indeed impact the inferences we are interested in. For instance, we do observe some differences in free choice effects with indefinites as compared to disjunctions (cf. (5))[4], which

[4]Thanks to an anonymous LI reviewer for encouraging us to discuss these data.

become less mysterious once the variability of domain restriction is factored in.

(5) a. John can present an article. But not Article 3.
b. John can present Article 1, Article 2, or Article 3. # But not Article 3.

In (5), it looks as if the free choice inference is absent or weaker with the indefinite, but this is readily explained by noting that, quite generally, the domain can be flexibly adjusted with the indefinite, and not with the fully explicit disjunction. Concretely, in the examples above, Article 3 may be dropped from the implicit domain of the indefinite if this is made necessary by the second clause. Domain restriction may thus explain the difference in strength of free choice effects between indefinites and disjunctions.

But the challenge is to explain a difference between free choice inferences and distributive inferences *with indefinites* (i.e. presence/strength of the former and absence/weakness of the latter). A theory of domain restriction would presumably be based on pressures in favor of minimally/maximally large domains, or pressures to end up with a maximally weak/strong meaning. But such pressures would probably not help distinguish between the cases in (3) and (4): to the extent that domain restriction may weaken free choice inferences, it would also weaken distributive inferences. If such a theory of domain restriction could be formulated though, one could then go on to try to seek independent evidence showing that indeed domain restriction works differently in child language from adult language.

8.3.2 Q-spreading blocked by intervention effects

There may be a piece of adult grammar that blocks the distributive inferences with indefinites for adults that is missing both in children's grammar and in semanticists' theory of grammar. This piece of grammar would be missed by children for lack of evidence up to a certain age or for lack of the competence needed to deploy it. Let us see what this piece of grammar could be like.

Chierchia (2013) develops an account for the distribution of negative polarity items (NPIs), such as the word *any* in English. He proposes that NPIs activate domain alternatives, which are then exhaustified by a syntactically present exhaustivity operator *exh*. Simplifying a lot, the semantic import of *exh* is to negate the alternatives activated by its prejacent which are not entailed by this prejacent. According to the proposal, NPIs are licensed when the exhaustification does not result in a contradiction. For instance, if one considers a domain D with three apples, D = {Apple 1, Apple 2, Apple 3}, *any* is licensed in (6a) because all of the domain alternatives of (6a), of the form as in (6b), are entailed by (6a). However, this account faces a challenge. Consider the ungrammatical (7a). Its domain alternatives are as in (7b). Clearly, the exhaustification of (7a) does not result in a contradiction — how come then that the NPI is not licensed there? To address this challenge, Chierchia (2013) proposes that *every* creates intervention effects: in a configuration such as (7a), in which the relevant operators are in the syntactic configuration schematized in (8), the alternatives of the negative polarity item cannot be exhaustified for syntactic reasons (simplifying somewhat, the common syntactic features of the universal quantifier and the indefinite are argued by Chierchia 2013 to create a minimality violation for the exhaustivity operator). This, by assumption, makes the negative polarity item ungrammatical in the scope of *every*.

(6) a. Mary didn't take any apples.

b. Mary didn't take an apple in D' (with D'⊆ D).

(7) a. *Each of the three girls took any apples.
b. Each of the three girls took an apple in D' (with D'⊆ D).

(8) *[*exh* [every [any]]]

The absence of q-spreading with indefinites in adult language may be the same type of situation: the universal quantifier creates intervention for the exhaustification of the alternatives of indefinites in the same position as *any* above, and therefore implicatures for sentences such as (4a) would be blocked. However, as already mentioned, adults do have q-spreading with disjunctions (i.e. distributive inferences, cf. Appendix 8.5.2). To the extent that this phenomenon results from exhaustification, this exhaustification happens within the same syntactic configuration as in (8), with the disjunction occupying the position of *any*. Hence, the universal quantifier is not intervening for all elements which activate alternatives and which are in its scope: it is not intervening for the exhaustification of disjunction, and this selectivity of intervention would have to be explained.[5]

If the implicature account of q-spreading in child language is correct and indeed adults do not exhibit q-spreading with indefinites due to intervention effects, then the difference between children and adults would be that children are not sensitive to intervention effects. It seems to us that this account involves many non-trivial pieces. But it has the virtue of being testable: children might accept polarity items in the scope of the universal quantifier, at least to the same extent that they show q-spreading effects.

8.4 Conclusion

Given current assumptions on the similarity of alternatives activated by indefinites and disjunctions, the fact that disjunctions, but not indefinites, trigger domain implicatures in the scope of the universal quantifier is puzzling. We have proposed that even if current semantic theories turn out to be incomplete for adults, they may be entirely correct for children, then, with q-spreading effects revealing the expected presence of domain implicatures. We discussed essentially two possibilities for the difference between children and adults: (i) children and semanticists are both wrong to assume that indefinites activate the same alternatives as disjunctions, (ii) indefinites do activate the same alternatives as disjunctions in child language, but children are missing a piece of adult grammar (perhaps flexible domain restriction of indefinites or sensitivity to intervention effects) that blocks some of their effects.

8.5 Appendix

In three experiments we explore adults' distributive inferences with disjunction. We first demonstrate that replacing the indefinite with a disjunction leads to an increase in rejections of universally quantified sentences in q-spreading-like contexts, which validates our methodology (Experiments 1 and 2). We then proceed to show that when the experimental context increases the salience of the subject, distributive inferences of disjunction decrease (Experiment 3), which

[5] Perhaps the universal quantifier intervenes only when the element which activates alternatives is of some type, surely *any* and indefinites are more similar with one another than they are to disjunctions. But note that items like *some* do trigger implicatures when in the position of *any* in configurations such as (8), so one would need to explain a broader pattern: the universal quantifier intervenes with *any* and indefinites, but not disjunctions and *some*.

is an effect similar to the one that Philip (2011) and Hollebrandse (2004) have observed with q-spreading with indefinites in children. Full data and analyses are available online at https://tinyurl.com/ya8w4gyo.

8.5.1 Experiment 1

First, we demonstrate that adults do not show q-spreading effects with indefinites. Participants judged q-spreading sensitive sentences such as (4a) as 'True', 'False', 'Neither'[6] with respect to three following types of images (corresponding to three conditions). FALSE: the sentence was false on the surface scope reading (some girl did not take any apple), TRUE: the sentence was true on the surface scope reading (all girls took an apple and all apples were taken), TARGET: the sentence was true on the surface scope reading (all girl took an apple, but crucially some apple was not taken (q-spreading)). Note that the inverse scope reading of (4a) was false across conditions. We varied the number of apples (2 or 3 for FALSE, 3 or 4 for TRUE, and 4 or 5 for TARGET), whether or not there was a girl who took more than one apple, and apple color (red and green), for a total of 36 items. In the target condition, 'Neither' or 'False' responses (to a larger extent than in other conditions) would be indicative of distributive inferences.

45 participants were recruited on Amazon Mechanical Turk, 3 of them were excluded from the analysis for giving more than 25% incorrect responses in TRUE and FALSE conditions. Generalized mixed effects linear regression model were fitted to the responses for TRUE and TARGET conditions with and without the condition as a fixed effect, and random by-participant intercepts and slopes (the maximal random effect structure for which convergence was achieved). The comparison did not reveal a significant effect of condition: $\chi^2(1) = 0.08, p = .76$ (cf. Figure 8.2, indefinite (left-most) panel): no evidence for q-spreading effects were found with indefinites in adults.

8.5.2 Experiment 2

Second, we show that simply replacing indefinites with disjunctions reveal q-spreading-like effects with adults. The experiment was thus similar except that the indefinite in the sentence (cf. (4a)) was replaced with a disjunction, as in (2a). As in Experiment 1, we were interested in the surface scope interpretation of (2a): none of the conditions validated the inverse scope reading.

46 other participants were recruited on Amazon Mechanical Turk, 4 were excluded from the analysis for their low scores on controls, as before. Mixed-models comparisons revealed a significant interaction between condition (TARGET vs TRUE) and experiment (indefinite vs disjunction): $\chi^2(1) = 11.0, p < .001$, thus showing that disjunctions make distributive inferences/q-spreading effects observable.

8.5.3 Experiment 3

Finally, we evaluate whether distributive inferences are sensitive to similar contextual manipulations as q-spreading effects with children (e.g., Crain et al. 1996; Philip 2011, a.o.). Very much following Hollebrandse 2004; Philip 2011 who found that the q-spreading is reduced with children when the topic is about the noun in the restrictor of the universal quantifier, we manipulated the relative salience of the subject and object, with the idea that putting salience away

[6] 'Neither' was a response option to accommodate people who derive distributive inferences, but would not go so far as to say that the sentence is false because of them. Overall, there were very few 'Neither' responses. They were collapsed with 'False' responses for the analysis.

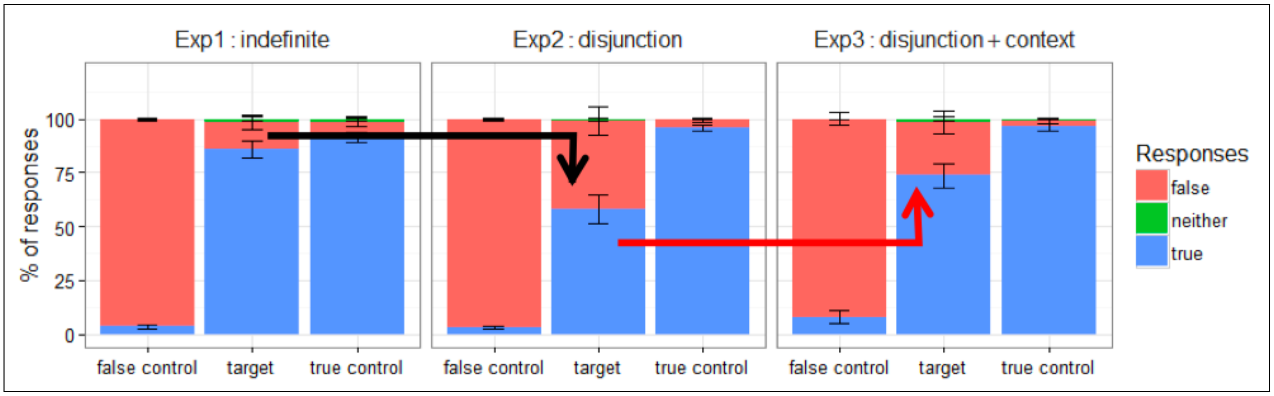


Figure 8.2: Distributive inferences can be measured as the proportion of false responses to the target condition (or more precisely as the increase of such responses compared to the baseline rate observed in adjacent true controls). The Figure represents adults' responses to universal sentences with an indefinite, with a disjunction, with a disjunction in the presence of additional contextual manipulation. Distributive inferences are absent in Exp. 1, they appear in Exp. 2 (black arrow), and they are also present in Exp. 3, but to a lesser extent (red arrow).

from the object should reduce the effect of its alternatives, and hence reduce q-spreading/distributive inferences. We did so with the following changes from Experiment 2: (i) the object was kept constant (i.e. boring) by keeping the same number of apples (=5) across trials and accordingly the same linguistic disjunction in the sentence, (ii) the subject was made more varied: the noun would now be either *girl* or *boy* and, in new filler sentences, the quantifier could be *exactly two*, *not every*, and *more than two*. The main changes were thus that there were (irrelevant) items in Experiment 3, for a total of 96 items (with only 24 items with the universal quantifier in subject position). As in Experiments 1 and 2, we were interested in the surface scope interpretation of (2a): none of the conditions validated the inverse scope reading of the universally quantified sentences.

45 participants were recruited on Mechanical Turk to do the task, 6 excluded using the same criterion as above. Mixed-models comparisons revealed a significant interaction between Condition (TARGET vs TRUE) and (i) Experiment (1 vs 3): $\chi^2(1) = 26.7, p < .001$, but also (ii) Experiment (2 vs 3): $\chi^2(1) = 4.40, p = .036$, thus showing that (i) distributive inferences/q-spreading effects exist but (ii) are weaker with the current contextual manipulation (red arrow in Figure 8.2).

Given that distributive inferences are instances of implicatures, that their rate of derivation is sensitive to contextual manipulations is not surprising. But these results open the way for a more substantial claim: a similar mechanism may be behind both the effect observed in Experiment 3 and the reduction of q-spreading observed in experiments with children: the lower salience of the the alternative trigger, which is disjunctions for adults and the indefinite noun phrase for children.

Ackwnoledgement
We would like to thank Amir Anvari, Moshe E. Bar-Lev, Gennaro Chierchia, Danny Fox, Martin Hackl, Naomi Havron, Daniel Hoek, Manuel Križ, Jeremy Kuhn, Mora Maldonado, Philippe Schlenker, Florian Schwarz, Jesse Snedeker, Andres Soria Ruiz, Benjamin Spector, Yasutada Sudo, Lyn Tieu, and the audience at SALT 28 for helpful discussion, critiques, and questions. The research leading to this work

was supported by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n.313610, and by ANR-17-EURE-0017.

Bibliography

Aloni, M. (2007). Free choice, modals, and imperatives. *Natural Language Semantics 15*(1), 65–94.

Anvari, A. (2018). Logical integrity. In *Semantics and Linguistic Theory*, Volume 28, pp. 711–726.

Aravind, A., J. de Villiers, P. de Villiers, C. J. Lonigan, B. M. Phillips, J. Clancy, S. H. Landry, P. R. Swank, M. Assel, H. B. Taylor, et al. (2017). Children's quantification with every over time. *Glossa: a journal of general linguistics 2*(1).

Bar-Lev, M. E. and D. Fox (2017). Universal free choice and innocent inclusion. SALT 2017.

Barker, C. (2010). Free choice permission as resource-sensitive reasoning. *Semantics and Pragmatics 3*, 1–38.

Barker, C. (2018). Negative polarity as scope marking. *Linguistics and Philosophy 41*(5), 483–510.

Barner, D. and A. Bachrach (2010). Inference and exact numerical representation in early language development. *Cognitive psychology 60*(1), 40–62.

Barner, D., N. Brooks, and A. Bale (2011). Accessing the unsaid: The role of scalar alternatives in childrenâĂŹs pragmatic inference. *Cognition 118*(1), 84–93.

Barwise, J. and R. Cooper (1981). Generalized quantifiers and natural language. In *Philosophy, language, and artificial intelligence*, pp. 241–301. Springer.

Beck, S. (2006). Intervention effects follow from focus interpretation*. *Natural Language Semantics 14*(1), 1–56.

Bergen, L., R. Levy, and N. Goodman (2016). Pragmatic reasoning through semantic inference. *Semantics and Pragmatics 9*.

Bott, L., T. M. Bailey, and D. Grodner (2012). Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language 66*(1), 123–142.

Bott, L. and I. A. Noveck (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language 51*(3), 437–457.

Brooks, P. J. and I. Sekerina (2006). Shortcuts to quantifier interpretation in children and adults. *Language Acquisition 13*(3), 177–206.

Buccola, B. and A. Haida (2018). Obligatory irrelevance and the computation of ignorance inferences. *Manuscript*.

Chemla, E. (2009). Similarity: Towards a unified account of scalar implicatures, free choice permission and presupposition projection. *Under revision for Semantics and Pragmatics.*

Chemla, E., C. Cummins, and R. Singh (2017). Training and timing local scalar enrichments under global pragmatic pressures. *Journal of Semantics 34*(1), 107–126.

Chemla, E., V. Homer, and D. Rothschild (2011). Modularity and intuitions in formal semantics: the case of polarity items. *Linguistics and Philosophy 34*(6), 537–570.

Chemla, E. and J. Romoli (2015). The role of probability in the accessibility of scalar inferences.

Chemla, E. and B. Spector (2011). Experimental evidence for embedded scalar implicatures. *Journal of semantics*, 359–400.

Chierchia, G. (1984). *Topics in the Syntax and Semantics of Infinitives and Gerunds.* Ph. D. thesis, University of Massachusetts at Amherst Ph. D. dissertation.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belleti (Ed.), *Structures and Beyond*, Volume 3, pp. 39–103. Oxford University Press.

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the "logicality" of language. *Linguistic inquiry 37*(4), 535–590.

Chierchia, G. (2013). *Logic in grammar: polarity, free choice and intervention.* Oxford University Press.

Chierchia, G., S. Crain, M. T. Guasti, A. Gualmini, L. Meroni, et al. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In *Proceedings of the 25th Boston University conference on language development*, pp. 157–168.

Chierchia, G., D. Fox, and B. Spector (2012). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In C. M. Klaus von Heusinger and P. Portner (Eds.), *Semantics: An international handbook of natural language meaning.* New York, NY: Mouton de Gruyter.

Crain, S. and R. Thornton (1998). *Investigations in universal grammar.* Cambridge, MA: MIT Press.

Crain, S., R. Thornton, C. Boster, L. Conway, D. Lillo-Martin, and E. Woodams (1996). Quantification without qualification. *Language Acquisition 5*(2), 83–153.

Cremers, A. and E. Chemla (2014). Direct and indirect scalar implicatures share the same processing signature. In S. Pistoia Reda (Ed.), *Pragmatics, Semantics and the Case of Scalar Implicatures*, Palgrave Studies in Pragmatics, Language and Cognition, pp. 201–227. London: Palgrave Macmillan.

Crnič, L. (2014). Non-monotonicity in NPI licensing. *Natural Language Semantics 22*(2), 169–217.

Crnič, L. (2014). Non-monotonicity in npi licensing. *Natural Language Semantics 22*(2), 169–217.

Crnič, L., E. Chemla, and D. Fox (2015). Scalar implicatures of embedded disjunction. *Natural Language Semantics 23*(4), 271–305.

De Neys, W. and W. Schaeken (2007). When people are more logical under cognitive load: Dual task impact on scalar implicature. *Experimental psychology 54*(2), 128–133.

Degen, J., M. H. Tessler, and N. D. Goodman (2015). Wonky worlds: Listeners revise world knowledge when utterances are odd. In *CogSci*.

Denić, M. (2015). Minimizing scope ambiguity hypothesis. In *ESSLLI 2015 Student Session*, pp. 54–65.

Denić, M., E. Chemla, and L. Tieu (2018). Intervention effects in npi licensing: a quantitative assessment of the scalar implicature explanation. *Glossa: a journal of general linguistics 3*(1), 49.

Dowty, D. (1979). Word meaning and montague grammar. dordrecht, holland: D.

Drenhaus, H., P. beim Graben, D. Saddy, and S. Frisch (2006). Diagnosis and repair of negative polarity constructions in the light of symbolic resonance analysis. *Brain and language 96*(3), 255–268.

Drenhaus, H., D. Saddy, and S. Frisch (2005). Processing negative polarity items: When negation comes through the backdoor. In S. Kepser and M. Reis (Eds.), *Linguistic evidence: Empirical, theoretical, and computational perspectives*, pp. 145–165. Berlin: Mouton de Gruyter.

Drenhaus, Heiner, Blaszczak, Joanna, and Schütte, Julianne (2007). Some psycholinguistic comments on npi licensing. In *Proceedings of Sinn und Bedeutung*, Volume 11, pp. 180–193.

Drozd, K. (2001). Children's weak interpretations of universally quantified questions. In M. Bowerman and S. Levinson (Eds.), *Language acquisition and conceptual development*, pp. 340–376. Cambridge: Cambridge University Press.

Drozd, K. and E. van Loosbroek (1998). Weak quantification, plausible dissent, and the development of children's pragmatic competence. In *Proceedings of the 23rd Boston University Conference on Language Development*, pp. 184–195.

Fauconnier, G. (1975a). Polarity and the scale principle. *Chicago Linguistics Society 11*, 107–126.

Fauconnier, G. (1975b). Polarity and the scale principle. *Chicago Linguistics Society 11*, 188–199.

Foppolo, F., M. T. Guasti, and G. Chierchia (2012). Scalar implicatures in child language: Give children a chance. *Language learning and development 8*(4), 365–394.

Fox, D. (2000). *Economy and semantic interpretation*, Volume 35. MIT press.

Fox, D. (2007). Free choice and the theory of scalar implicatures. In *Presupposition and implicature in compositional semantics*, pp. 71–120. Springer.

Fox, D. (2014). Cancelling the maxim of quantity: Another challenge for a gricean theory of scalar implicatures.

Fox, D. and M. Hackl (2006). The universal density of measurement. *Linguistics and Philosophy 29*, 537–586.

Fox, D. and R. Katzir (2011). On the characterization of alternatives. *Natural language semantics 19*(1), 87–107.

Fox, D. and R. Katzir (2019). Modularity and iterated rationality models of scalar implicatures.

Franke, M. (2011). Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics and Pragmatics 4*, 1–82.

Freeman, N. H., C. G. Sinha, and J. A. Stedmon (1982). All the cars–which cars? from word meaning to discourse analysis. *Children thinking through language*, 52–74.

Gajewski, J. R. (2002). On Analyticity in Natural Language. Ms., MIT.

Gajewski, J. R. (2005). *Neg-raising: Polarity and presupposition.* Ph. D. thesis, Massachusetts Institute of Technology.

Gazdar, G. (1979). *Pragmatics, implicature, presuposition and logical form.* Academic Press.

Geurts, B. (2003). Quantifying kids. *Language acquisition 11*(4), 197–218.

Geurts, B. and F. van Der Slik (2005). Monotonicity and processing load. *Journal of semantics 22*(1), 97–117.

Gibson, E. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown.* Ph. D. thesis, School of Computer Science, Carnegie Mellon University.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition 68*(1), 1–76.

Goodman, N. D. and A. Stuhlmüller (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in cognitive science 5*(1), 173–184.

Gotzner, N. and J. Romoli (2017). The scalar inferences of strong scalar terms under negative quantifiers and constraints on the theory of alternatives. *Journal of Semantics 35*(1), 95–126.

Grice, H. P. (1975). Logic and conversation. *Syntax and Semantics*, 41–58.

Guasti, M. T., G. Chierchia, S. Crain, F. Foppolo, A. Gualmini, and L. Meroni (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and cognitive processes 20*(5), 667–696.

Guerzoni, E. (2006). Intervention effects on NPIs and feature movement: Towards a unified account of intervention. *Natural Language Semantics 14*(4), 359–398.

Herburger, E. and S. Mauck (2007). A new look at ladusaw?s puzzle. In H. Zeijlstra and J.-P. Soehn (Eds.), *Workshop on Negation and Polarity*, pp. 78–84.

Hollebrandse, B. (2004). Topichood and quantification in L1 dutch. *International Review of Applied Linguistics and Language Teaching 2*, 203–215.

Homer, V. (2008). Disruption of NPI linensing: The case of presuppositions. In *Proceedings of SALT 18*.

Homer, V. (2012). Domains of polarity items.

Horn, L. R. (1972). *On the semantic properties of logical operators in English.* Ph. D. thesis, University of California, Los Angeles.

Huang, Y. T., E. Spelke, and J. Snedeker (2013). What exactly do numbers mean? *Language Learning and Development 9*(2), 105–129.

Inhelder, B. and J. Piaget (1964). *The early growth of logic in the child: Classification and seriation.* Routledge and Kegan Paul.

Kadmon, N. and F. Landman (1993a). Any. *Linguistics and Philosophy 4.2*, 279–298.

Kadmon, N. and F. Landman (1993b). Any. *Linguistics and philosophy 16*(4), 353–422.

Kamp, H. (1973). Free choice permission. In *Proceedings of the Aristotelian Society*, Volume 74, pp. 57–74.

Kang, H.-K. (2001). Quantifier spreading: Linguistic and pragmatic considerations. *Lingua 111*(8), 591–627.

Katzir, R. (2007). Structurally-defined alternatives. *Linguistics and Philosophy 30*(6), 669–690.

Katzir, R. (2014). On the roles of markedness and contradiction in the use of alternatives. In *Pragmatics, semantics and the case of scalar implicatures*, pp. 40–71. Springer.

Klinedinst, N. W. (2007). Plurals, possibilities, and conjunctive disjunction. *UCL Working Papers in Linguistics 19*, 261–284.

Krifka, M. (1991). Some remarks on polarity items. *Semantic universals and universal semantics*, 150–189.

Krifka, M. (1995). The semantics and pragmatics of polarity items. *Linguistic analysis 25*(3-4), 209–257.

Kroch, A. S. (1974). *The semantics of scope in English.* Ph. D. thesis, Massachusetts Institute of Technology.

Ladusaw, A. W. (1979). *Polarity sensitivity as inherent scope relations.* Ph. D. thesis, University of Texas at Austin.

Ladusaw, W. (1986). Principles of semantic filtering. In *Papers from the 5th Western Conference on Formal Linguistics, Stanford Linguistics Association, Stanford, Ca.*

Lahiri, U. (1998). Focus and negative polarity in Hindi. *Natural Language Semantics 6*, 57–123.

Linebarger, M. C. (1987). Negative polarity and grammatical representation. *Linguistics and Philosophy 10*(3), 325–387.

Magri, G. (2009a). A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural language semantics 17*(3), 245–297.

Magri, G. (2009b). *A theory of individual-level predicates based on blind mandatory scalar implicatures: constraint promotion for Optimality Theory.* Ph. D. thesis, Ph. D. dissertation, MIT.

Marty, P. and E. Chemla (2013). Scalar implicatures: working memory and a comparison with 'only'. *Frontiers in Psychology 4*(403).

Mascarenhas, S. (2014). *Formal Semantics and the Psychology of Reasoning: Building new bridges and investigating interactions.* Ph. D. thesis, New York University, New York.

Menendez-Benito, P. (2005). *The grammar of choice.* Ph. D. thesis, University of Massachusetts Amherst.

Meyer, M.-C. (2013). *Ignorance and grammar.* Ph. D. thesis, Massachusetts Institute of Technology.

Meyer, M.-C. (2014). Deriving hurford's constraint. In *Semantics and Linguistic Theory*, Volume 24, pp. 577–596.

Muller, H. and C. Phillips (2018). *Negative polarity illusions*, Volume Oxford Handbook of Negation. Oxford University Press.

Nouwen, R. (2010). Two kinds of modified numerals. *Semantics and Pragmatics 3*, 3–1.

Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition 78*(2), 165–188.

Noveck, I. A. and A. Posada (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language 85*(2), 203–210.

Pagliarini, E., C. Bill, J. Romoli, L. Tieu, and S. Crain (2018). On childrenâĂŹs variable success with scalar inferences: Insights from disjunction in the scope of a universal quantifier. *Cognition 178*, 178–192.

Papafragou, A. and J. Musolino (2003). Scalar implicatures: experiments at the semantics–pragmatics interface. *Cognition 86*(3), 253–282.

Parker, D. and C. Phillips (2016). Negative polarity illusions and the format of hierarchical encodings in memory. *Cognition 157*, 321–339.

Philip, W. (1995). *Event quantification in the acquisition of universal quantification.* Ph. D. thesis, University of Massachusets, Amherst.

Philip, W. (2011). Acquiring knowledge of universal quantification. In J. de Villiers and T. Roeper (Eds.), *Handbook of generative approaches to language acquisition. Studies in Theoretical Psycholinguistics*, Volume 41, pp. 351–394. Dordrecht: Springer.

Philip, W. and M. Takahashi (1991). Quantifier spreading in the acquisition of every. In T. Maxfield and B. Plunckett (Eds.), *Papers in the Acquisition of WH: Proceedings of the UMass Roundtable, University of Massachusetts Occasional Papers*, pp. 283–301. Amherst, MA.

Progovac, L. (2000). Negative and positive feature checking and the distribution of polarity items. *Negation in Slavic*, 88–114.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Roeper, T. (2007). *The prism of grammar: How child language illuminates humanism.* Boston MA:Bradford Books.

Roeper, T. and J. de Villiers (1991). The emergence of bound variable structures. *University of Massachusetts Occasional Papers: Papers in the Acquisition of WH*, 267–282.

Roeper, T., B. Z. Pearson, and M. Grace (2011). Quantifier spreading is not distributive. In *Proceedings of the 35th Boston University Conference on Language Development*, pp. 526–539. Somerville, MA: Cascadilla Press.

Romoli, J. and F. Schwarz (2015). An experimental comparison between presuppositions and indirect scalar implicatures. In F. Schwarz (Ed.), *Experimental perspectives on presuppositions*, pp. 215–240. Springer.

Rothschild, D. (2006). Non-monotonic NPI-licensing, definite descriptions, and grammaticalized implicatures. In *Proceedings of SALT*, Volume 16, pp. 228–240.

Rullmann, H. (1995). *Maximality in the semantics of wh-constructions.* Ph. D. thesis, University of Massachusetts at Amherst.

Saddy, D., H. Drenhaus, and S. Frisch (2004). Processing polarity items: Contrastive licensing costs. *Brain and language 90*(1-3), 495–502.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy 27*(3), 367–391.

Schulz, K. and R. van Rooij (2006). Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy 29*(2), 205.

Shao, J. and H. Neville (1998). Analyzing semantic processing using event-related potentials. *Newsletter for the Center for Research in Language 11*(5), 3–20.

Singh, R., K. Wexler, A. Astle, D. Kamawar, and D. Fox (2016). Children interpret disjunction as conjunction: consequences for the theory of scalar implicatures. *Natural Language Semantics.*

Spector, B. (2006). *Aspects de la pragmatique des opérateurs logiques.* Ph. D. thesis, Université de Paris 7.

Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures. In U. Sauerland and P. Stateva (Eds.), *Presuppositions and Implicatures in Compositional Semantics.* Palgrave-Macmillan.

Spector, B. (2014a). Global positive polarity items and obligatory exhaustivity. *Semantics and Pragmatics 7*, 11–1.

Spector, B. (2014b). Scalar implicatures, blindness and common knowledge: Comments on magri (2011). In *Pragmatics, semantics and the case of scalar implicatures*, pp. 146–169. Springer.

Spector, B. (2018). What do we learn from game-theoretic pragmatics? SALT28 presentation.

Spector, B. (2019). Modified numerals. In R. Z. Matthewson, Meier (Ed.), *Wiley Companion to Semantics.*

Sprouse, J., M. Wagers, and C. Phillips (2012). A test of the relation between working memory capacity and syntactic island effects. *Language 88*(1), 82–123.

Stalnaker, R. (2002). Common ground. *Linguistics and philosophy 25*(5-6), 701–721.

Steinhauer, K., J. E. Drury, P. Portner, M. Walenski, and M. T. Ullman (2010). Syntax, concepts, and logic in the temporal dynamics of language comprehension: Evidence from event-related potentials. *Neuropsychologia 48*(6), 1525–1542.

Szabolcsi, A. (2004). Positive polarity–negative polarity. *Natural Language & Linguistic Theory 22*(2), 409–452.

Szabolcsi, A., L. Bott, and B. McElree (2008). The effect of negative polarity items on inference verification. *Journal of Semantics 25*, 411–450.

Takahashi, M. (1991). Children's interpretation of sentences containing every. In T. Maxfield and B. Plunckett (Eds.), *Papers in the Acquisition of WH: Proceedings of the UMass Roundtable, University of Massachusetts Occasional Papers*, pp. 303–329. Amherst, MA.

Tieu, L. and J. Lidz (2016). Npi licensing and beyond: Children's knowledge of the semantics of any. *Language Acquisition 23*(4), 311–332.

Tieu, L., J. Romoli, P. Zhou, and S. Crain (2015). Children's knowledge of free choice inferences and scalar implicatures. *Journal of Semantics 33*(2), 269–298.

Tieu, L., J. Romoli, P. Zhou, and S. Crain (2016). Children's knowledge of free choice inferences and scalar implicatures. *Journal of Semantics 33*(2), 269–298.

Tieu, L., K. Yatsushiro, A. Cremers, J. Romoli, U. Sauerland, and E. Chemla (2017). On the role of alternatives in the acquisition of simple and complex disjunctions in french and japanese. *Journal of Semantics 34*(1), 127–152.

van der Wouden, T. (1997). *Negative Contexts.* Routledge Studies in Germanic Linguistics.

van Rooij, R. and K. Schulz (2004). Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information 13*(4), 491–519.

Vasishth, S., S. Brüssow, R. L. Lewis, and H. Drenhaus (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science 32*(4), 685–712.

Von Fintel, K. (1993). Exceptive constructions. *Natural Language Semantics 1*(2), 123–148.

von Fintel, K. (1994). *Restrictions on quantifier domains.* Ph. D. thesis, University of Massachusetts, Amherst.

Westerståhl, D. (1989). Quantifiers in formal and natural languages. In *Handbook of philosophical logic*, pp. 1–131. Springer.

Xiang, M., B. Dillon, and C. Phillips (2009). Illusory licensing effects across dependency types: Erp evidence. *Brain and Language 108*(1), 40–55.

Xiang, M., J. Grove, and A. Giannakidou (2013). Dependency-dependent interference: Npi interference, agreement attraction, and global pragmatic inferences. *Frontiers in psychology 4*, 708.

Yanilmaz, A. and J. E. Drury (2018). Prospective npi licensing and intrusion in turkish. *Language, Cognition and Neuroscience 33*(1), 111–138.

Yurchenko, A., D.-B. Den Ouden, J. Hoeksema, O. Dragoy, J. C. Hoeks, and L. A. Stowe (2013). Processing polarity: Erp evidence for differences between positive and negative polarity. *Neuropsychologia 51*(1), 132–141.

Zwarts, F. (1998). Three types of polarity. *Plurality and Quantification 69*, 177–238.

RÉSUMÉ

Il a été démontré que plusieurs phénomènes linguistiques corrèlent avec certaines propriétés logiques de la phrase dans laquelle ils se trouvent, telles que les implications logiques de la phrase, ou les relations logiques entre la phrase et certaines phrases alternatives. Dans cette thèse, nous explorons (A) la question d'un éventuel rôle causal de ces propriétés logiques vis-à-vis de tels phénomènes linguistiques, et (B) la question de la manière dont ces propriétés logiques sont calculées. Il y a dans l'ensemble deux options concernant (B) : (i) ces propriétés logiques pourraient être calculées par un système formel n'ayant pas accès aux connaissances générales, appelons ce système formel *la grammaire* (Fox et Hackl 2006, entre autres), ou bien (ii) ces propriétés logiques pourraient être calculées de manière post-grammaticale. Nous abordons ces questions dans deux cas : la légitimation des éléments à polarité négative et la dérivation des implicatures scalaires. Nos conclusions s'inscrivent dans la lignée d'un grand nombre de résultats attestant d'un rôle causal de propriétés logiques au sein de ces deux phénomènes. Par ailleurs, nos résultats suggèrent que les propriétés logiques corrélées aux NPIs font intervenir des calculs post-grammaticaux, tandis que les propriétés logiques intervenant dans les implicatures scalaires sont calculées au niveau grammatical.

MOTS CLÉS

langage, logique, éléments à polarité négatives, implicatures scalaires, sémantique expérimentale

ABSTRACT

Numerous linguistic phenomena have been shown to correlate with some logical properties of the sentence with which they occur. Examples of such properties are logical entailments supported by the sentence, or the logical relation between it and some alternative sentence. In this thesis, we explore (A) whether these logical correlates play a causal role in the linguistic phenomena in question, and (B) at what level these computations of logical correlates are performed. There are two broad possibilities concerning (B): (i) these logical correlates could be computed in a formal system that does not have access to contextual knowledge, call this system grammar (Fox and Hackl 2006, a.o.), or (ii) they could be computed post-grammatically, and therefore have access to contextual knowledge. We investigate these questions in relation to the licensing of negative polarity items (NPIs) and the derivation of scalar implicatures. Our findings add to a growing body of evidence that logical correlates of NPI licensing and scalar implicatures play a causal role in these phenomena. Furthermore, our results suggest that logical correlates to NPI licensing are calculated post-grammatically, while logical correlates to scalar implicatures are calculated grammar-internally.

KEYWORDS

language, logic, negative polarity items, scalar implicatures, experimental semantics