PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'École Normale Supérieure

Facteurs structurels et contextuels en pragmatique

Soutenue par
Émile ENGUEHARD
Le 4 octobre 2021

École doctorale n° 158
Cerveau, cognition, comportement

Spécialité
Sciences cognitives

Composition du jury :

Emmanuel CHEMLA
Directeur de recherche, ENS/EHESS/CNRS
Président

Maria ALONI
Professeure associée (UHD), Université d'Amsterdam
Rapportrice

Maribel ROMERO
Professeure, Université de Constance
Rapportrice

Uli SAUERLAND
Professeur, Leibniz-ZAS
Examinateur

Yasutada SUDO
Professeur associé, University College London
Examinateur

Benjamin SPECTOR
Directeur de recherche, ENS/EHESS/CNRS
Directeur de thèse

ENS DEC DÉPARTEMENT D'ÉTUDES COGNITIVES

On the role of alternatives at the semantics-pragmatics interface

Structural and contextual factors in pragmatics

Émile Enguehard

Ph.D. thesis prepared under the supervision of
Benjamin Spector

Institut Jean Nicod
École Normale Supérieure

Résumé Cette thèse a pour thème fondamental le rôle des alternatives dans la sémantique et la pragmatique des langues naturelles, et tout particulièrement la question de la division du travail exacte entre la sémantique et la pragmatique pour ce qui touche aux phénomènes sensibles aux alternatives.

On compte parmi les plus importants de ces phénomènes celui des implicatures scalaires et de l'exhaustification, qui fait l'objet des deux premiers chapitres. Le chapitre 2 (après une introduction) concerne le potentiel des numéraux modifiés comparatifs, les expressions comme "plus de trois", à engendrer des implicatures scalaires. Je montre que contrairement à ce qui a été avancé dans la littérature, on peut observer des implicatures scalaires dans les phrases faisant intervenir ces expressions, et que leur présence dépend de facteurs liés à la connaissance générale ainsi que la rotondité du nombre. Je propose une théorie basée sur une exhaustification obligatoire avec des échelles lexicales, et où le choix de l'échelle dépend d'une questions en discussion (QUD) déterminée pragmatiquement. Le chapitre 3 est consacré à un problème bien connu des théories de l'exhaustification qui la place dans le système sémantique : ces théories permettent l'exhaustification enchassée, mais doivent contraindre sa distribution pour être en accord avec les observations. L'argument du chapitre est qu'il est possible de dériver cette distribution à partir d'une contrainte globale de connexité sur le sens. Cette contrainte est inspirée de travaux sur le caractère naturel ou non des concepts, et est liée de façon plausible à un biais cognitif très général.

Dans le chapitre 4, il est question d'une famille d'exemples modaux, les constructions de suffisance minimale, dont le sens tel qu'observé s'écarte de ce que prédisent les idées généralement admises sur leurs constituants. Je compare plusieurs approches de ce problème inspirés de la littérature, qui font toutes intervenir des échelles pragmatiques ainsi qu'une forme de sensibilité aux alternatives, mais en la situant dans des composantes différentes de la phrase. La conclusion est qu'aucune des ces approches n'est tout à fait satisfaisante.

Le chapitre 5 se veut une illustration de l'usage des modèles probabilistes de la pragmatique, qui ont été proposés comme alternative à l'exhaustification, pour dériver des résultats formels d'intérêt pour la linguistique. La question dont il s'agit est le problème posé par Horn sur la lexicalisation des opérateurs logiques formant ce qu'on appelle le carré d'Aristote : pourquoi est-ce qu'un certain lexique (A, E, I) est très souvent observé alors qu'un autre (A, E, O), tout aussi expressif, ne l'est pas ? Il est montré que l'optimalité du lexique attesté se dérive d'une généralisation naturelle sur le sens des mots de contenu à l'aide d'un modèle de l'usage du langage fondé sur la théorie de la décision.

Pour finir, les deux derniers chapitres portent sur la sémantique des questions. Je montre que certaines généralisations sur la projection des présuppositions dans les questions polaires coordonnées sont problématique pour l'approche des questions fondée sur la sémantique alternative (aussi appelée Hamblin-Karttunen), ainsi que pour d'autres théories existantes, et ce parce que les théories importantes de la projection des présupposition ne permettent pas de dériver les données. Je propose comme solution l'adoption d'une sémantique des questions plus riche que celle des alternatives, qui établisse un parallélisme plus clair entre les questions polaires et les déclaratives, et je présente deux théories de ce type, l'une fondée sur la sémantique inquisitive et la logique trivalent, dans le chapitre 6, et l'autre fondée sur la sémantique dynamique, dans le chapitre 7.

Abstract The overarching theme of this thesis is the role of alternatives in natural language semantics and pragmatics, and particularly the question of the precise division of labour between semantics and pragmatics when it comes to alternative-sensitive phenomena.

Prominent among these phenomena is that of scalar implicatures and exhaustification, which the first two chapters focus on. Chapter 2 (after the introduction) discusses the potential for triggering scalar implicatures of comparative modified numerals, expressions such as "more than three". I find that contrary to earlier claims, scalar implicatures are observed in sentences where these expressions occur, and that whether they are possible or not depends on certain world-knowledge factors as well as the roundness of the number. I propose an account based on obligatory exhaustification with lexical scales, where the choice of scale is based on a pragmatically-determined Question Under Discussion. Chapter 3 addresses a well-known problem for theories of exhaustification that locate it in the semantics: these theories allow for embedded exhaustification, but need to constrain its distribution to fit observations. The main claim developed in the chapter is that we can derive that distribution from a whole-meaning connectedness constraint. This constraint is inspired by work on the naturalness of concepts and can plausibly be related to a very general cognitive bias.

In Chapter 4, I discuss a family of modal examples called Minimal Sufficiency constructions whose observed meaning is at odds with what received views of the meaning of their component parts predict. I contrast a variety of accounts of these examples inspired from the literature, all based on pragmatic scales as well as some form of alternative-sensitivity, though locating it in different components of the sentence, and show that none of them are fully adequate. Chapter 5 is a case study in the use of probabilistic models of pragmatics, which have been proposed as an alternative to exhaustification, for the derivation of formal results of interests to linguistics. The problem of interest is Horn's proposed universal concerning the lexicalization of the logical operators forming the so-called Square of Aristotle, where a certain lexicon (A, E, I) is widely attested while another, equally expressive one (A, E, O) is not. It is shown that the optimality of the attested lexicon can be derived from a natural generalization on the meaning of content words within a decision-theoretic model of language use, without recourse to lexical or cognitive markedness as explanatory factors.

Finally, the last two chapters are concerned with question semantics. I show that certain patterns of presupposition projection in coordinated polar questions present a challenge to the alternative semantics (a.k.a. Hamblin-Karttunen) view of questions, as well as various other existing theories, in that established theories of presupposition projection fail to derive the data. I submit that we need semantics for questions that go beyond alternatives and feature greater parallelism between polar questions and declaratives, and I propose two such accounts, one based on inquisitive semantics and trivalent logic in Chapter 6, and one based on dynamic semantics in Chapter 7.

Acknowledgements

The writing of this thesis is the conclusion of five years I spent at DEC, the cognitive science department of ENS, which I came to as an informal master's student before embarking on a Ph.D. These years were overall a very pleasant experience, thanks in particular to Benjamin Spector, my supervisor, not just because he was an extremely helpful mentor but also because he was the one who welcomed and guided me in when I initially decided to investigate the strange field of linguistics. I would also like to thank the other members of our group, LINGUAE, who helped make these years a formative experience for me, in particular Emmanuel Chemla, Salvador Mascarenhas, Jeremy Kuhn, and Philippe Schlenker, as well as the other Parisian researchers who steered me in my first forays into linguistics, in particular Giorgio Magri, Tal Linzen, and Heather Burnett.

My heartfelt thanks also go to those who made me the honor of being part of my defense committee, Maria Aloni, Maribel Romero, Emmanuel Chemla again, Uli Sauerland, and Yasutada Sudo.

Five months of this period were spent at MIT as part of a visitorship; for this time I would like to thank all the members of MIT Linguistics, who were as welcoming to me as I could hope for; in particular my thanks go to Roger Schwarzschild, who sponsored my visit, as well as Patrick Elliott, Keny Chatain, Enrico Flor, Filipe Hisao Kobayashi, Vincent Rouillard, and Stanislao Zompì.

Of course and above all, my time at DEC would have been a lot less interesting had I not made friends there, and for this reason I would like to thank all those with whom I had the repeated pleasure of exchanging ideas of all sorts, emotions, and of course drinks. My very special thanks go to (in rough order of appearance) my esteemed teacher Milica Denić, Mathias Sablé-Meyer, Mora Maldonado, Amir Anvari, Andrés Soria Ruiz, Chloé Tahar, Jeremy Kuhn again, Salvador Mascarenhas again, Rachel Dudley, Brian Buccola, Diego Feinmann, Nadine Bade, Nur Lan, WooJin Chung, and Victor Carranza.

Finally, I thank Louise. This second time we go through thesis-writing together should be the last, not that there is anything we can't handle.

Contents

Chapter 1

Introduction: alternatives in semantics and pragmatics

1.1 On this thesis

This thesis gathers together a number of articles, proceedings papers and manuscripts written over the course of my doctorate, with each piece forming one chapter. While these papers are not part of a unique narrative, they do share a thematic relationship: each of them is a stab, from a different angle, at the question of how alternatives contribute to the determination of meaning in natural language. The present chapter seeks to introduce and contextualize these pieces, locate them within the field of formal semantics and pragmatics and relate them to the various puzzles of the field related to the notion of alternatives.

1.2 Comparative modified numerals as a case study of scalar implicature

Comparative modified numerals and scalar implicatures Chapter 2 is a study of the pragmatics of comparative modifed numerals, phrases such as *more than seven*. In (1), we see that these phrases can trigger specific pragmatic inferences: upon hearing this example, we tend to infer that Mary walked some distance between 7 and 8kms. Yet, there is an intuition that, if Mary in fact walked 10kms, (1) is still true. This suggests that the inference that Mary walked less than 8kms is pragmatic in nature, rather than being part of the semantic content.

(1) Mary walked more than seven kilometres.
⇝ Mary walked between seven and eight kilometres.

The pragmatic inference seen in (1) can be straightforwardly analysed as a scalar implicature: if Mary had in fact walked more than 8kms, then the speaker might have said (2) truthfully, and they would have been more informative because (2) logically entails (1) — it describes a smaller, more precise set of situations. Assuming that the speaker is trying to be both truthful and maximally informative, they would never say (1) in this situation. Taking the contrapositive, if they did say (1), then Mary did not walk more than 8kms.

(2) Mary walked more than 8kms.

This form of reasoning is known as Gricean reasoning, because it is based on Grice's (1991) maxims of Quality (the speaker is being truthful) and Quantity (the speaker is being maximally informative). (2) plays the role of an alternative, and the resulting inference is an implicature.

Alternatives, Horn scales, and exhaustification The result of Gricean reasoning is crucially dependent on the choice of the set of alternatives, the sentences we compare the utterance to. Indeed, if we assume that (3) is an alternative to (1), based on the same reasoning, we derive an implicature to the effect that John walked more than 8kms. Such an implicature is in fact never observed; if it were, together with the earlier one, we would reach a contradiction. In such a situation, (2) and (3) are called symmetric alternatives. Here, it has to be the case that (2) always counts as an alternative to (1) for the purposes of Gricean reasoning, while (3) never does; the theoretical challenge raised by this fact is called the symmetry problem.[1] One standard solution is to assume that alternatives originate from Horn scales, collections of logically ordered elements. In particular, the discrete numeral scale $\langle 1, 2, 3, \ldots, 7, 8, \ldots \rangle$ will make (2) an alternative to (1).

(3) John walked between 7 and 8kms.

In other cases, symmetric alternatives are desirable: (4) implies that the speaker is unsure whether John walked more than 7kms, or whether he walked exactly 7kms. This can be derived through Gricean reasoning if we assume that both sentences in (5) constitute (symmetric) alternatives: had the speaker be certain that one of these two sentences were true, they would have said it instead.[2] We can derive these alternatives from the modifier scales ⟨at least, exactly⟩ and ⟨at least, more than⟩.

(4) John walked at least 7kms
$\rightsquigarrow$ The speaker is unsure whether John walked exactly 7kms, or more than that.

(5) a. John walked exactly 7kms.
b. John walked more than 7kms.

Given a theory of how implicatures arise from alternative sets,[3] we can investigate what the alternative set should be for a given sentence, so that the appropriate implicatures are generated. This will in turn inform theories of implicature generation.

The alternatives of comparative modified numerals Modified numerals are a particularly fertile ground for this sort of investigation, given the clear logical structure of the domain. In chapter 2, I propose that the alternatives of comparative modified numerals are

[1] See Breheny et al. 2018 for a survey of issues related to the symmetry problem. One consequence of note is that alternative sets need to have a certain shape in order to make non-trivial predictions; in particular, they should not be closed under logical operations.

[2] Here we see that earlier we have been assuming that the speaker knew exactly which situation we were in; without this assumption, we only derive primary implicatures ("The speaker does not hold the belief that X is the case") rather than secondary implicatures ("The speaker holds the belief that X is not the case").

[3] An example of such a theory is Sauerland's (2004) procedure. We can also define an *exhaustification operator* that computes secondary implicatures: if φ is a proposition, corresponding to an utterance s, and A is a set of propositions, then $\textsc{exh}(\varphi; A) = \varphi^+$ where φ^+ is the enriched meaning of s when the alternatives are given by A. See Spector 2016 for a discussion of the precise semantics of EXH.

While exhaustification operators are ususally defined in such a way that only secondary implicatures are generated, it has been proposed to include primary implicatures as well (Buccola and Haida 2019, a.o.).

based on the discrete numeral scale, as well as a modifier scale that relates comparative numerals to bare numerals (interpreted under an "at least" reading): $\langle n, \text{more than } n\rangle$. The assumption that the alternatives of modified numerals are discrete goes against Fox and Hackl's (2006) *Universal Density of Measurement* proposal, which relies on a continuous scale. This is because Fox and Hackl seek to derive an absence of implicatures, based on examples like (6) where indeed no implicature is observed. However, the possibility of an implicature in (1), which is a novel observation, as well as in (7), as noted by Cummins, Sauerland, and Solt (2012), is consistent with the existence of discrete scales. In fact, we need not just the basic numeral scale, but also "granular" discrete scales based on roundness such as $\langle 10, 20, 30, \ldots\rangle$ etc.

(6) There are more than 7 people here.

(7) There are more than 20 people here.
$\rightsquigarrow$ There are fewer than 30 people.

Under one of two proposals discussed in chapter 2, the absence of an implicature in (6) is due to the fact that the enriched meaning would be "there are more than 7 people but fewer than 8", which is nonsensical. This follows from the assumption that exhaustification is blind to contextual entailment patterns, as proposed by Magri (2009); blind exhaustification conspires with world-knowledge elements to make it so that exhaustified interpretations of sentences using comparative modified numerals are only possible if round numbers are used or if the quantity being discussed is dense. This proposal is contrasted with another one based on non-blind contextual entailment, and I discuss predictions made by the assumptions on Horn scales when it comes to embedded environments.

The role of relevance In connection with Grice's maxim of Relation, the alternatives that enter Gricean calculations are thought to be subject to a relevance constraint. Formalisations of this tend to involve a notion of Question under Discussion (QUD, after Roberts 1996), a possibly implicit question that the utterance is interpreted as an answer to. Indeed, asking an explicit question affects implicatures; the implicature of (1) is observed when it is an answer to a quantity question, but not when it is an answer to a polar question, as seen in (8). This suggests that (2) is relevant to the former question but not the latter; this is indeed what the notion of relevance defined by Groenendijk and Stokhof (1984) predicts.[4]

(8) a. Q: How much did John walk?
A: John walked more than 7kms.
$\rightsquigarrow$ John did not walk more than 8kms.

b. Q: Did John walk more than 7kms?
A: John walked more than 7kms.
$\not\rightsquigarrow$ John did not walk more than 8kms.

I propose in chapter 2 that exhaustification is obligatory once the QUD is determined. Along with the structural theory whereby (6)'s exhaustified reading is nonsensical, this pre-

[4]Importantly, relevance constraints do not help towards solving the symmetry problem. Indeed, relevance should intuitively be closed under the usual logical operations (negation, conjunction, disjunction — for instance, if p is a relevant proposition, then $\neg p$ is relevant too), which would guarantee that all alternatives have a symmetric counterpart, and thus that no secondary implicature is ever generated.The notion of relevance formalized by Groenendijk and Stokhof (1984) is indeed closed under logical operations. This is why there needs to be other constraints on the alternative set, such as those set by Horn scales.

dicts that (6) cannot be a felicitous answer to a *how many* question. Thus, upon hearing (6), speakers will infer that the QUD is a polar question similar to (8b), which makes all alternatives irrelevant. We should therefore observe an irrelevance inference from (6): the inference that the number of people does not matter. I defend in chapter 2 that this prediction is correct, based in part on observations due to Buccola and Haida (2017). An associated prediction is that there should never be a scalar implicature from (6), which explains the observations Fox and Hackl (2006) base their theory on.

Interim conclusion A high-level outlook on chapter 2 is that it provides a case study of how, given certain assumptions on exhaustification procedures and on what is relevant in a given discourse context, we can infer the appropriate alternative set for a given utterance; then, based on the theory of Horn scales, we can then make further predictions for other sentences.

The account reduces the role of pragmatics in the theory: there is no actual optionality for implicature generation and it is only the selection of a particular QUD that is subject to pragmatic considerations. Arguably, the presence of multiple granularity scales postulated in chapter 2 is also subject to pragmatic selection, given that a round numeral like 20 may be interpreted either within the unit scale or the tens' scale. One could choose to unify these two mechanisms, and make granularity a matter of QUD: there is a single scale of all numerals, but only alternatives using sufficiently round numerals are relevant.

Either way, the significance of roundness is interesting in that it is not clear whether it is more of a conceptual notion or more of a linguistic one, which connects to the question of whether alternatives are determined in terms of conceptual or linguistic representations.

1.3 Minimal sufficiency: pragmatic scales and alternative sensitivity

The prejacent problem, and related issues The focus of chapter 4 is a category of examples involving pragmatic or logical scales and license a specific set of inferences called Minimal Sufficiency. The puzzle that these examples raise is that they involve necessity modals, but they do not license the inferences that common wisdom about necessity would lead us to expect; what they express is rather weaker. This is shown below.

(9) To get good cheese, you only have to go to the North End [a part of Boston].
$\not\rightsquigarrow$ You can't find good cheese anywhere else.
(von Fintel and Iatridou 2007)

(10) To qualify, you have to get a silver medal.
$\not\rightsquigarrow$ If you get a gold medal, you won't qualify.

(11) To be over the limit, you have to drink TWO beers.
$\not\rightsquigarrow$ If you drink no beer, but you drink two glasses of wine, you won't be over the limit.

Von Fintel and Iatridou (2007) call the case of (9) the prejacent problem, because they observe that (12), the prejacent of (9), does not receive a similar weak reading. The analysis they develop is based on a novel, weak semantics for *only*. I argue that the examples should be analysed in parallel, and therefore that the puzzle we are faced with is not about *only*.

(12) To get good cheese, you have to go to the North End.
⇝ You can't find good cheese anywhere else.

Pragmatic scales and scalar operators The meaning of all these examples can be stated in a similar way: we infer that the embedded condition is *sufficient* to meet the goal, and that doing less will not do; but we do not infer that doing more or doing something entirely unrelated will not work. I call this set of inferences "Minimal Sufficiency". Crucially, they involve a notion of ranked alternatives — hence "less" and "more". In (11) it is the numeral scale, while in (10) it is the (highly conventional) scale of medals (⟨bronze, silver, gold⟩) and in (9) it is an *ad hoc* scale of locations, perhaps ranked by effort to get there.

Sensitivity to this kind of "pragmatic scales" is a known property of scalar operators such as *only* (Klinedinst 2004). For instance, what (13) means is that John does not have a "better" degree than a bachelor's, for instance a master's degree. To analyse examples such as (13), we might have to distinguish between "scalar *only*" and "logical *only*", where logical *only* is based on entailment, and scalar *only* is based on a pragmatic ordering such as the one in (14).

(13) John only has a bachelor's degree.

(14) high school diploma $\prec$ bachelor's $\prec$ master's

Given that scalar *only* is very reminiscent of logical *only*, there has been proposals to the effect that only scalar *only* exists, and logical *only* is just a special case where the ordering coincides with logical entailment (e.g. Klinedinst 2004) or that only logical *only* exists, and pragmatic scales correspond to non-literal readings of certain expressions such that they are in fact ordered by entailment (Magri 2017; Alonso-Ovalle and Hirsch 2018).

Chapter 4 discusses, in addition to the proposal of von Fintel and Iatridou (2007), two potential towards a solution to the challenge raised by Minimal Sufficiency examples. The first one follows Alonso-Ovalle and Hirsch (2018) in reducing pragmatic scales to logical ones, which is implemented in the syntax through the use of a covert operator AT-LEAST. The resulting weakening is, it turns out, not sufficient, and I propose to move to a weaker operator AT-LEAST-DEG, though this is also problematic in some respects. The other route consists in providing an account of the modal as a alternative-sensitive operator, that is, an operator whose semantics involves an alternative set, like *only*. Proposals along this line are offered by by Villalta (2008) and Krasikova (2010). I show however that this does not lead to a satisfying solution.

Connection to general considerations on alternatives The discussion offered in chapter 4 is inconclusive, but it bears on several important issues regarding alternatives. One question is how pragmatic scales arise. Under all analyses discussed, the difference between (9) and (15), which does not easily receive a Minimal Sufficiency reading, is that the presence of *only* forces a reading of the embedded proposition as being part of a pragmatic scale. Thus, pragmatic scales are not lexically determined. Within the first analytic approach discussed above, I propose that pragmatic scales in fact correspond to cases where a salient mapping between alternatives and a gradable property — corresponding perhaps to some notion of "impressiveness" — is available.

The potential for an alternative-sensitive analysis of modals that would be aware of pragmatic scales also raises the issue of the integration of scalarity with theory of alternatives. Beyond scalar implicatures, which we already discussed, one of the main contexts in which alternatives have been discussed is that of alternative-sensitive (or focus-sensitive) operators, including *only*. These operators are sensitive to the placement of prosodic prominence, and the associated patterns can be derived in a theory assuming alternatives are generated by replacing focussed elements, and only focussed elements, with elements of the same semantic category or "domain alternatives" (Rooth 1992). Fox and Katzir (2011) argue that the alternative sets involved in scalar implicatures and focus-sensitivity should be analysed the same way, among other things because the symmetry problem also occurs for *only*; they propose that alternatives are generated roughly along the lines of focus semantics, but with an added constraint that replacement candidates have to be at least as simple as the original expression in some formal sense.[5,6] A complete unified theory of alternatives should of course make a place for pragmatic scales; they can be detected both when it comes to alternative-sensitivity, as chapter 4 and the above discussion show, and in the case of scalar implicatures, as seen in (15).

(15) If John has a bachelor's degree, he will find a job.
⇝ If John has only a high school diploma, he might not find a job.

Yet another issue raised by the matter of Minimal Sufficiency has to do with the division of labour between semantics and pragmatics. The first route based on AT-LEAST-DEG predicts an extremely weak meaning for the examples; in particular, it does not imply, in the case of (9), that going to the North End is in fact a way to get good cheese. I suggest that this might not be a problem, and that the missing inference might be a manner implicature, but we lack the formal theory that predicts its existence. On the contrary, the second route is consistent with the view that focus-sensitive processes are generally located in the grammar (which we will come back to in the next section).

Finally, we observe in both (10) and (11) that exhaustification is obligatory: these sentences are not true if you in fact have to get a gold medal or to drink three beers. This lends support to the idea, already put forward in the previous section, that exhaustification is in general obligatory and appearance to the contrary is due to the indeterminacy of alternatives. If we focus on the scalar reading, which is associated to a specific set of alternatives (the pragmatic scale), then we cannot avoid the inferences due to exhaustification.

1.4 The grammatical theory of exhaustification and its difficulties

As we have seen, the procedure through which scalar implicatures are computed can be formally summarized through an operator EXH. While the precise definition one should give to EXH is debated (see Spector 2016), its effect is closely related to the semantics of

[5]This constraint is argued by Fox and Katzir (2011) to solve the symmetry problem: for instance, (3) is not an alternative to (1) because it is more complex. We will see, however, that the case of indirect implicatures remains problematic. The constraint can be thought of as a formalization of Grice's maxim of Manner.

[6]Fox and Katzir's theory predicts that only alternatives obtained through replacing focused elements trigger implicatures. This generalization admits some counterexamples, such as the following:

(i) Mary DID meet some of the professors.
⇝ Mary did not meet all of the professors.

It is remarkable that this is not what we observe with comparative modified numerals: an implicature is really impossible in (8b).

only, giving rise to the idea that EXH is a "silent *only*".[7]

While EXH can be viewed as a pure formal device standing in for Gricean reasoning, it has been proposed that there is in fact such an operator in the grammar, which, much like *only*, can occur in embedded positions. Such a theory predicts the possibility of embedded implicatures: if a sentence S denotes proposition φ under traditional assumptions, and can trigger an implicature to the effect that proposition ψ is true, then we expect to find a more complex sentence where S occurs as an embedded clause denoting $\varphi \land \psi$. For instance, we expect that (16a) could correspond to a Logical Form similar to (16b),[8] which would yield a reading paraphrasable as (16c).

(16) a. Every student did some of the homework.
b. [Every student] [1 [EXH [t_1 did some of the homework]]]
c. Every student did some but not all of the homework.

Chierchia, Fox, and Spector (2011) argue that this prediction is correct, as do Chemla and Spector (2011) on the basis of experimental evidence. An interesting aspect of the *grammatical theory* of implicatures, as it is called, is that it locates a phenomenon generally considered to be pragmatic in the semantic system. Under the grammatical theory, scalar implicatures are in fact part of the truth conditions of sentences. Our judgment that they can be falsified without falsifying the sentence is due either to the possibility of a parse of the sentence with no occurrence of EXH, or to the fact that the alternative set is determined pragmatically. Pragmatic reasoning in this picture serves solely to disambiguate when needed, and does not generate inferences directly. This grammaticalization of Gricean reasoning has the advantage that it helps explain cases where exhaustification appears to be obligatory, and does away with the problem that straightforward application of Grice's maxims only derives primary implicatures (The speaker does not consider that X is the case) and not secondary implicatures (The speaker considers that X is not the case), requiring the stipulation of an extra "Epistemic Step" (cf. for instance Sauerland 2004).

One of the main difficulties the grammatical theory faces is that embedded implicatures are only observed in a limited subset of possible embedded environments. For instance, when *some* occurs in a negative environment, as in (17a), there could in principle be an embedded implicature, and there is no obvious reason why the resulting reading (paraphrased in (17b) for (17a)) should be deviant. Yet, inasmuch as this reading is observed at all, it requires a very specific prosody, with strong emphasis on *some*, and an echoic context.

(17) a. I doubt that Mary did some of the homework.
b. I expect that Mary did all or none of the homework.

In order to rule out embedded implicatures in (17a) and various other environments, there needs to be a constraint on the distribution of EXH. Chierchia, Fox, and Spector (2011) propose that the choice of parse follows the Strongest Meaning Hypothesis (SMH): parses that lead to logically stronger interpretations are systematically preferred. Roughly speaking, this theory predicts that monotonicity is the crucial property governing the distribution of embedded EXH; it should be observed in upwards-entailing environments, but not

[7] Beyond the issue of scalar implicatures, the notion of exhaustification has also been used to analyse the interpretation of questions and answers. In fact, it is in this context that Groenendijk and Stokhof (1984) introduce the notation EXH, though the operator they define has a different type. See also section 1.6.

[8] Here I assume that EXH is a propositional operator and I resort to binding to make it take scope under the quantifier. This is merely for the sake of illustration; one could resort to another scope-taking mechanism, or simply generalize EXH in such a way that it can take predicative arguments.

in downwards-entailing environments. This correctly predicts a dispreference for embedded implicatures in (17a) and a variant of this theory has been argued in detail by Fox and Spector (2018) to deliver good predictions in general.

Chapter 3 is a contribution to the debate on the distribution of embedded exhaustification. Together with Emmanuel Chemla, we propose an alternative to accounts based on monotonicty, in the form of a *connectedness* (or *convexity*) constraint. The idea is that parses that result in *non-connected* meanings are dispreferred, using a notion of "non-connected meanings" that more or less corresponds, intuitively, to a certain kind of disjunctive propositions, including that expressed by (17b). We show in particular that our proposal correctly rules out embedded implicatures in certain upwards-entailing environments — specifically, below existential quantifiers, and within a member of a disjunction — and propose that it is related to a more general cognitive bias that has been argued by Gärdenfors (2004) to affect conceptual representations.

Beyond the matter of embedded exhaustification, a novel proposal made in Chapter 3 is that the concept of connectedness can be applied to propositional meanings. The method through which we operationalize this idea makes crucial use of alternatives:[9] we assume that possible worlds are ordered through the subset of alternatives that are true at each world. Thus, if the alternatives are "John did some of the homework" and "John did all of the homework", a world where John did some but not all of the homework, so that one alternative is true, is "below" a world where John did all of the homework, where both alternatives are true, and "above" a world where John did nothing, where no alternative is true. From this order we derive a notion of *in-betweenness*, and we define connectedness based on that notion. Under this view, the alternative set is what structures the space of possibilities, in a way that is reflected in cognitive representations. Note that for the order not to be trivial, we need the alternatives to be logically related and of the same monotonicity; in other words they need to resemble the denotation of a question in answer set semantics, rather than be closed under negation or be partition cells (cf. section 1.6). Interestingly, this corresponds to the alternative sets that theories of exhaustification need to assume in order not to derive trivial results.

1.5 The role of probabilistic models in linguistic theorizing

The Square of Aristotle Chapter 5, written with Benjamin Spector, is concerned with Horn's puzzle concerning the so-called square of Aristotle (or square of opposition) of logical operators (Horn 1973). The square of opposition consists in 4 categories of quantified statements, *A*, *E*, *I* and *O*, exemplifed in (18), and Horn's observation is that across languages, the *O* corner is almost never associated to a lexical quantifying determiner — there is no **nall* or **nevery*.

(18) a. *A*: All cats are black.
b. *E*: No cat is black.
c. *I*: Some cats are black.
d. *O*: Not all / *nall cats are black.

[9] Most of the discussion in Chapter 3, however, considers connectedness at an intuitive level and does not crucially depend on the choice of operationalization.

We know that, thanks to scalar implicatures, just three of the determiners, either $\{A, E, I\}$, or $\{A, E, O\}$, are sufficient to distinguish between the three possible states of things (all cats are black, some but not all are, or none are). The puzzle consists in explaining why we always observe $\{A, E, I\}$ in practice. The solution proposed by Horn and more recently by Katzir and Singh (2013) is that O and negative (or perhaps: downward-entailing) operators in general are dispreferred at the cognitive level. Part of the evidence for this fact, which we already discussed in the context of indirect implicatures above, is that negative operators are syntactically or morphologically more complex, which makes the explanation potentially circular. The explanation we propose, in contrast, is based on general considerations on language use, derived from game-theoretic models of pragmatics.

Game-theoretic models Game-theoretic models of pragmatics, such as the Rational Speech Act model (RSA, Goodman and Stuhlmüller 2013) or the Iterated Best Response model (IBR, Franke 2011, a.o.) offer a radical alternative to exaustification-based approaches of scalar implicatures.

In these models, speakers and listeners are assumed to be drawing their inferences through an iterative signalling game. We model speaker and listener behaviours as conditional probability distributions: a listener picks out worlds given utterances, while a speaker picks out utterances given worlds. Thus we write as $L(w|u)$ the probability that a listener L who heard utterance u assigns to world w, and as $S(u|w)$ the probability that a speaker utters u in w. The purpose of both participants is to maximize the chance the listener picks out the actual world in reaction to the message, that is, to maximize $L(w_0|u_0)$ where w_0 is the actual world (unknown to the listener) and u_0 is the actual utterance.

The basic listener L_0 interprets utterances literally, ruling out situations that are incompatible with the truth, and sticking to its prior probabilities otherwise. Based on a model of listener behaviour L_n, we derive a speaker S_{n+1} who is maximizing the objective under the assumption that the listener follows L_n, as well as a listener L_{n+1} who is assuming that the speaker follows S_{n+1}. In this way, recursively, we can define two infinite sequences of models $(S_n)_n$ and $(L_n)_n$. Real-world speakers and listeners are assumed to be described by some higher-order model (say S_5, L_5) or by the limit of the sequence in infinity, if it converges.

The exact recursion rule will vary from one implementation to another. In IBR, the (simplified) general idea is that speakers pick out the message that directly maximizes the objective, while listeners perform a Bayesian update based on speaker behaviour, which depends on a prior on worlds p_0, before picking out the *a posteriori* most probable world:

(19) IBR recursion:
  a. $S_n(u|w) = \mathbb{1}[u = \arg\max_{u'} L_{n-1}(w|u')]$
  b. $L_n(w|u) = \mathbb{1}[w = \arg\max_{w'} S_n(u|w')p_0(w')]$

In RSA, both sides are probabilistic. Listeners perform a pure Bayesian update, while speakers maximize utility in a noisy way based on a so-called soft-max rule:

(20) RSA recursion:
  a. $S_n(u|w) \propto \exp(\lambda \log L_{n-1}(w|u))$
  b. $L_n(w|u) \propto S_n(u|w)p_0(w)$

where $\lambda > 1$ is a constant called the rationality parameter. In both models, we can add a message cost term c to the speaker's objective, to model the fact that the speaker might prefer simpler messages; for instance in RSA this is done through the intermediary of a utility function U:

(21) RSA speaker with cost: $S_n(u|w) \propto \exp(\lambda U(w, u; L_{n-1}))$
where: $U(w, u; L) = \log L(w|u) - c_u$

It is interesting to note that the speaker decision rule in these models can be thought of as an implementation of the Gricean maxims. Because the basic listener L_0 takes only semantic truth into account, S_1 and all higher-order speakers will never utter false messages, and thus they follow Quality. The maximization of $L(w|u)$ can be thought of as Quantity, given that it consists in getting the listener to assign a minimal probability to worlds other than the actual one (a quantity given by $1 - L(w|u)$). Finally, the incorporation of a cost term corresponds to Manner. Relation is reflected in the stipulation of a limited alternative set, as in more traditional accounts. The main innovation is therefore in the amalgamation of Quantity and Manner in a single optimization step, rather than having them enter the procedure in different ways.

Indirect implicatures Game-theoretic models are able to deal with the symmetry problem in much the same way that the theory of Fox and Katzir (2011) does: due to the incorporation of a cost term, it suffices to assume that the offending alternative as higher cost to derive the desired implicature. What is particularly interesting is that these models are also capable of dealing with the case of indirect implicatures, that is, implicatures originating from negative sentences, which prove problematic for the structural complexity theory.

Going back to our earlier example of comparative modified numerals, consider (22). In much the same way as the basic example (23) (repeated from (1)), (22) triggers an implicature to the effect that John walked between 7 and 8kms. As mentioned in chapter 2, this can be analysed in an entirely symmetric way to the earlier case, simply assuming that the Horn scales associated to *fewer than* numerals are the symmetric counterparts to those associated to *more than* numerals.

(22) John walked less than 8kms.
$\rightsquigarrow$ John walked more than 7kms.

(23) John walked more than 7 kms.

As discussed by Breheny et al. (2018), the theory of Fox and Katzir (2011) does not in fact predict this symmetry. In order to derive the implicature of (23), we need to assume that (22) is not an alternative; otherwise (22) and (24) (repeated from (2)) would be symmetric alternatives. This is not problematic, because there are reasons to think that (22) is more complex than (23) and (24). First, *less than* numerals can lead to split scope readings (Penka 2012), suggesting they are decomposable elements, and second, there is some evidence that monotone decreasing quantifiers, including *less than* numerals, are more difficult for speakers to process than monotone increasing quantifiers, including *more than* numerals (Geurts and van der Slik 2005).

(24) John walked more than 8kms.

However, when the roles are reversed and we are trying to derive the implicature of (22), there is no reason to exclude (23), which is simpler and obtained by replacement, from the alternatives. Since there is no reason to exclude (25), the alternative obtained by changing the numeral, either, we find ourselves with symmetric alternatives and predict, incorrectly, an ignorance inference.

(25) John walked less than 7kms.

The numerical decision rule of speakers in a game-theoretic model allows us to include both (23) and (22) within each other's alternatives without endangering the implicature in either case, while also making an interesting prediction regarding the alternation between the two. Without going into the details of the demonstration — a discussion of parallel examples is found in chapter 5, section 5.2.1 — we predict that (23) is more likely to be used when there is a high prior probability that John walked less than 7kms, and then the implicature obtains, whereas (22) is more likely to be used when there is a high prior probability that John walked more than 8kms, and then the implicature obtains. The predictions regarding priors are correct; if John is not very athletic and is generally reticent to walk more than 5kms, it is strange to utter (22), and it is not clear that the implicature obtains; however if John usually walks 10kms a day, the utterance is very natural and the implicature is clear. The alternation between *less than* and *more than* numerals is seen particularly clearly when talking about temperatures, as seen in the following examples:

(26) Context: *the outside temperature is 35C.*
Q: What's the temperature outside?
a. A: It's more than 30.
b. #A: It's less than 40.

(27) Context: *the outside temperature is -5C.*
Q: What's the temperature outside?
a. #A: It's more than -10.
b. A: It's less than 0.

In conclusion, game-theoretic models seem to allow us to relax our constraints on alternative sets without running into the symmetry problem, while also predicting certain correct patterns of sensitivity to priors.[10] This is made possible by the optimization step that unites Quantity and Manner and makes it possible to compare potentially symmetric alternatives without deriving degenerate results.

Application to Horn's puzzle and theoretical significance Probabilistic models — by which I mean both game-theoretic models based on iterated reasoning, and more generally any decision-theoretic model involving a notion of numerical utility — have been applied to a wide range of phenomena beyond scalar implicatures, such as the interpretation of gradable adjectives (Lassiter and Goodman 2017; Qing and Franke 2014), the effect of stress placement (Bergen and Goodman 2015), or irony (Kao et al. 2014). Nevertheless, they have often not been considered adequate replacement for formal-logical theories. One problem is the very sensitivity to priors that we just demonstrated; in many cases, it leads to problematic predictions.[11] A more general problem with the literature on game-theoretic

[10]This sensitivity is however problematic in some ways; see footnote 11. Additionally, the simplified game-theoretic models we present here are not equipped to deal with speaker ignorance, and will therefore not be adequate for situations where we do want symmetric alternatives.

[11]For instance, in RSA, in the situation where John is *a priori* likely not to have walked much, we predict that speakers are unlikely to use (22), which seems correct. We also predict, however, that if speakers do use (22), it will lead to an inference that John walked less than 7kms; this inference is, never observed. If we replace the soft-max decision rule with a perfect rationality rule (corresponding to taking $\lambda \to +\infty$ in RSA, or to the default in IBR), we predict that (22) is never spoken in this situation, which means that it will be uninterpretable for the listener. In the case of comparative modified numerals, it is not clear that the

models is that it mostly relies on numerical simulations. In general, authors will demonstrate that a given linguistic phenomenon can be reproduced in a certain model, but they will not explore the full range of predictions for different settings of the parameters (rationality, priors, costs, etc.). Thus, it is not possible to assess the generality of the predictions, nor the generative power of the theory. Similarly, because the results are derived through numerical simulation, it is difficult to extract any theorems or other analytical statement to which the predictions made by formal-logical theories can be compared. In this sense, most proposals are more of a proof of concept than a serious competitor for established theoretical accounts of the same phenomenon.[12]

The solution to Horn's puzzle we offer in chapter 5 can be seen as an example of how decision-theoretic models can be used to derive analytical results. This requires to keep one's ambitions contained and focus on a highly specific point.

Our solution relies crucially on the patterns of sensitivity to priors predicted by utility-based models which we discussed above.[13] Given that the speaker is most likely to use I in situations where A is particularly unlikely to be true, and most likely to use O in situations where E is particularly unlikely to be true, our argument is essentially that if the former sort of situations are more common than the latter, then we expect to find I to be lexicalized rather than O. Furthermore, we argue that the former situations are indeed more common, due to the fact that most predicates occurring in language denote a minority of things. This intuition is formally implemented through a meta-probabilistic model (where the probability distributions governing speaker behaviour are themselves random variables). Our argument demonstrates the possibility of novel analytic results from decision-theoretic models, at the price of high specificity — for instance, we cannot extend the comparison beyond the two lexica $\{A, E, I\}$ and $\{A, E, O\}$. The basic property of decision-theoretic models when applied to implicatures, which is that they conflate Quantity and Manner in a single optimization step, is crucial to our reasoning, as this is what allows direct comparison of I and O statements.

Other than the meta-theoretical point, our proposal offers a solution to Horn's puzzle that does not involve cognitive or linguistic representations; instead, it depends on a general property of lexicalization patterns in relation to the external world — the fact that predicates denote a minority of things — as well as general principles of communication — the form of the models, which as we have seen are an implementation of the Gricean maxims. Thus, it can be considered to be a true explanation for the dispreferrence towards downwards monotonicity that we observe in language. Interestingly, our solution to Horn's puzzle makes it a pragmatic phenomenon, brought about by speaker behaviour tendencies, which is in con-

---

prediction is bad, as (22) is indeed highly infelicitous in this situation and seems to prompt for a reevaluation of priors. However, when it comes to the more canonical example of the scale ⟨some, all⟩, even in cases of degraded felicity the implicature is reliably observed; this is shown in (i). In this instance, we would want our generation of both direct and indirect implicatures not to be affected by priors. Similar problems of sensitivity to priors are noted by Fox and Katzir (2020), focussing on IBR.

(i) Context: *children do not have the right to vote and rarely try to vote anyway.*

a. Some of the people who came to vote were children.
⇝ Not all of them were.

b. ? Not all of the people who came to vote were children.
⇝ Some of them were.

[12] Similar criticisms can be levelled at the literature on machine learning approaches to linguistics (see Baroni 2021 for discussion of this point).

[13] Specifically, we look at the behaviour of S_1 in a fully-rational variant of RSA.

trast with earlier proposals linking it to representational asymmetries and therefore making it a syntactic-semantic phenomenon.

1.6 Issues in question semantics

The final context in which we will discuss alternative sets is that of question semantics. Proposition sets have indeed also been introduced as a potential representation for questions, within the so-called answer set theory. As we are going to see, a lot of challenges raised by the composition and representation of questions echo those found in relation to the nature of alternative sets in declarative sentences. Chapters 6 and 7 argue for the need for a representation of questions that goes beyond mere alternatives, in particular so that we can define particular connectives.

The answer set theory The dominant paradigm in question semantics is the answer set theory, which sees questions as sets of propositions, corresponding intuitively to the possible answers. There are two families of systems, some based on Hamblin semantics (Hamblin 1976), and some based on the work of Karttunen (1977) — for this reason answer set theories are also referred to as Hamblin-Karttunen (or H/K) semantics. The differences do not really matter for our purposes and we can focus on Hamblin semantics. In Hamblin semantics, a question directly denotes the set of its possible answers, which is obtained by replacing the *wh*-word with elements of its domain. The typical analysis of polar questions, as well as two potential analyses for a basic *wh*-question are found below:[14]

(28) ⟦Did Mary come?⟧ = {“Mary came”, “Mary did not come”}

(29) a. ⟦Who came?⟧ = {“Mary came”, “John came”}
b. ⟦Who came?⟧ = {“Mary came”, “John came”, “Mary and John came”}

While many members of the answer set can be true at a given world, in general, only the most informative true answer is considered acceptable; this is seen in (30). There are, however, counter-examples to this pattern called “mention-some” questions (as opposed to “mention-all”); (31) is an example. In order to account for mention-some examples, we need to distinguish between positive and negative relevant information: (31d) is not an acceptable mention-some answer, which justifies the choice of only “positive” propositions as part of the answer set. Another justification is that, with the answer set as it is, the interpretation of an answer to a mention-all question such as (30c) can be analysed as the exhaustification of the answer, taking the entire answer set as the alternatives.[15] Indeed, as we already mentioned in section 1.3, (30c) is interpreted to mean that nobody else than Mary and John knows the password (and this explains why (30a) and (30b) are inappropriate: their enriched interpretation is false). If negative propositions were included, we would not be able to derive the exhaustive interpretation of answers in this way, as we would run into the symmetry problem.

(30) Context: *two people know the password, Mary and John.*
Q: Who knows the password?
a. #A: Mary knows the password.

[14]Here the sentences between quotation marks represent propositions, not linguistic objects.

[15]Fox (2018) argues that with the appropriate definition of exhaustification, this is true of mention-some questions as well.

b. #A: John knows the password.
c. A: Mary and John know the password.

(31) Context: *same as above.*
Q: Who can tell me the password?
a. A: Mary can tell you the password.
b. A: John can tell you the password.
c. A: Mary and/or John can tell you the password.
d. #A: Mary cannot tell you the password.

From the historical perspective of Hamblin semantics, the absence of negative propositions from the denotation does not seem remarkable, because in simple cases, it proceeds from the method of composition: we essentially replace the *wh*-word with type *e* entities and there is no opportunity for negative propositions to arise. In this respect, Hamblin semantics echo the alternative semantics for focus of Rooth (1992). Indeed, the relation between *wh*-words (including a potential covert *wh*-word in polar questions) and Hamblin question denotations is essentially the same as the relation between focussed elements and the alternative set in focus semantics. Yet the fact that, much like eligibility towards triggering an implicature, answerhood cannot be reduced to relevance (as it is not closed under logical operations), and that in both cases, an additional bias towards "positive" elements is observed, is not self-evident. Furthermore, recall that Fox and Katzir's (2011) argument that Gricean alternatives and focus alternatives should be unified relies on the fact that the symmetry problem can be observed for focus alternatives. To the point, Spector (2008) observes that higher-order questions where the *wh*-word ranges over quantifiers exist, and that in order to account properly for the semantics of these questions when embedded, the *wh*-word needs to range specifically over monotone increasing quantifiers. Thus, the same restriction to monotone increasing quantifiers is needed for the alternatives involved in scalar implicatures, focus sensitivity, and questions, suggesting that ultimately they constitute a single phenomenon.

Other theories of questions A variety of other approaches to question semantics exist. A prominent one is partition theory (Groenendijk and Stokhof 1984). In partition theory, a question is an equivalence relation over the set of possible worlds; two worlds are related if the question is answered in the same way in both of them. In this way, worlds are partitioned into cells, corresponding to answers. All this presupposes that there is only one good answer in each world; thus partition theory can make fewer distinctions between questions than answer set theory, where we tolerate that several answers might be true at a world.[16] This makes it impossible to deal with mention-some questions, as we have already seen.

Partition theory is important to our earlier discussion of exhaustification because Groenendijk and Stokhof (1984) use it to offer a formal definition of relevance: the basic principle is that a proposition is relevant if it can be written as a disjunction of cells. This notion of relevance is closed under the usual logical operations (negation, conjunction and disjunction), as, intuitively, it should be, and therefore, as we discussed, does not solve the symmetry problem.

Answer sets are, in a sense, a more expressive representation of questions than partitions. Indeed, we can obtain partition cells from an answer set denotation by exhaustifying

[16] In this respect partition theory is an implementation of the programme for question semantics set out by Hamblin (1958), while later work based on the compositional system presented in Hamblin 1976 is not.

each member while taking the rest of the set as alternatives. Thus, the benefits or partition theory, such as the possibility to define relevance and other logical concepts, are not compromised when we adopt answer set semantics.

Other theories of questions include categorial theories, where questions have a higher-order functional type (e.g. Krifka 2001; Xiang 2021), inquisitive semantics, which is close to alternative semantics but with more solid ties to formal logic (Ciardelli, Groenendijk, and Roelofsen 2013), as well as a variety of dynamic theories (e.g. Aloni and van Rooy 2002; Haida 2007; Dotlačil and Roelofsen 2019; Li 2019). Two novel theories, which we will come back to below, are also introduced in chapters 6 and 7. In much the same way that partition theory can be derived in a sense from answer set theories, all these theories generally stand in relation to one another. The diagram in Figure 1.1 represents an informal expressivity hierarchy; if a theory or family of theories is below another one, it can represent fewer distinctions.[17] For instance, all theories above answer set semantics can distinguish (29a) and (29b), an important distinction to make in plural semantics. Similarly, all theories above inquisitive semantics can represent mention-some answers (and are subject to the symmetry problem, in the same way that answer set theories are). To give a final example, all theories above trivalent inquisitive predicates are yes/no-asymmetric, that is, they readily distinguish (32a), (32b) and (32c); this is relevant to the fact that these questions differ in the biases they express and the ways they affect discourse (Bolinger 1978; Büring and Gunlogson 2000; Biezma and Rawlins 2012; Sudo 2013).[18] As we are going to see, it is also relevant to presupposition projection in coordinated questions.

(32) a. Is Syldavia a monarchy?
b. Is Syldavia not a monarchy?
c. Is Syldavia a monarchy or not?

Pragmatic effects in (coordinated) polar questions The discussion so far has focussed on the answerhood conditions of questions, and indirectly their embedded semantics. However, another avenue of investigation regarding the appropriate representation for questions is the pragmatic inferences that question themselves license. This is the focus of the final two chapters.

Chapter 6 is chiefly concerned with presupposition projection. Common wisdom is that in polar question such as (33), presuppositions triggered inside the nucleus project; in fact, this is often used as a test for presupposition status. Nevertheless, I observe that in both conjunctions and disjunctions of polar questions, presuppositions triggered inside the nucleus of the right-hand member can be filtered (using the terminology of Karttunen (1973)). This is shown in (34). The observed pattern is in fact exactly similar to what is observed by Karttunen (1973) for declarative sentences with the same structure.

(33) Is the Syldavian monarch a progressive?
$\rightsquigarrow$ Syldavia is a monarchy.

[17] It is important to note that this diagram represents nothing else than the expressivity of semantic representations. In particular, various theories can be linked to a formal logic of questions, with notions of entailment, connectives etc., and these logics will form a hierarchy of their own which will not be the one in the diagram.

[18] Note though that it is also possible to distinguish positive and negative questions by assuming extra structure is present; see Romero and Han 2004 for an example. Biezma and Rawlins (2012) also propose a yes/no-asymmetric treatment of matrix polar questions within answer set theory.

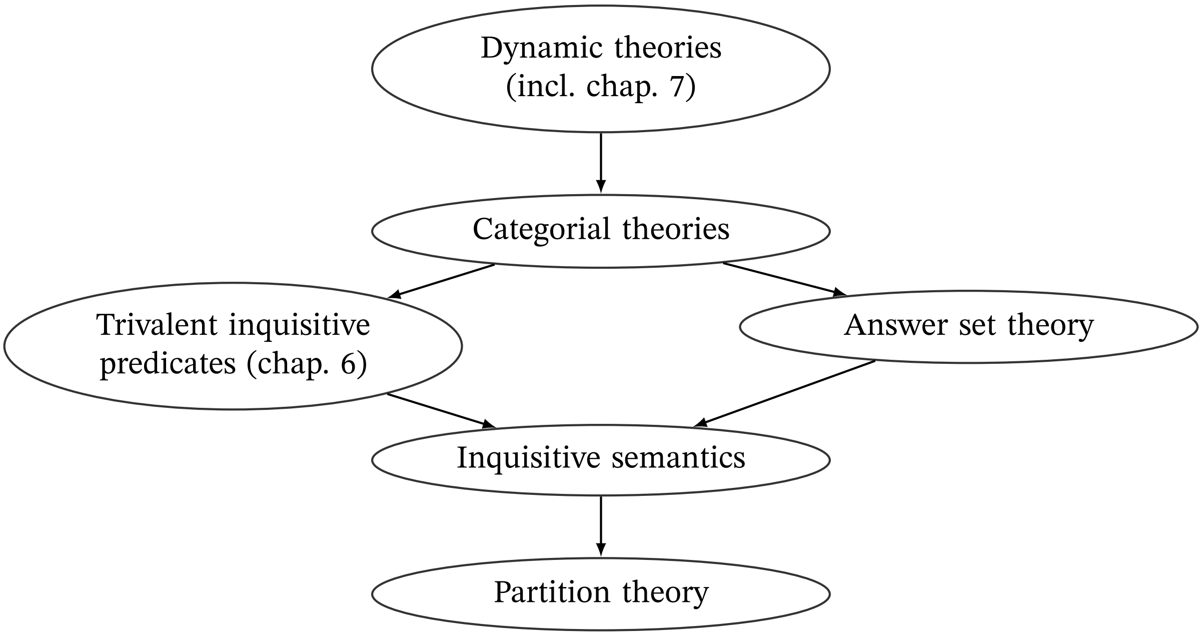

Figure 1.1 – Informal hierarchy of expressivity within theories of question semantics.

(34) a. Is Syldavia a monarchy and is the Syldavian monarch a progressive?
b. Is Syldavia a republic or is the Syldavian monarch a progressive?
⇸ Syldavia is a monarchy.

Presupposition projection in declaratives is a relatively well-understood topic, and a number of theories have been proposed that derive filtering patterns from the semantics of logical operators (e.g. Schlenker 2008). I show that it is difficult to reconcile these theories with existing views of question semantics. The problem is essentially that those theories that include an account of question coordination (the answer set theory, inquisitive semantics, the partition theory) also lack yes/no-asymmetry, which is essential to derive the yes/no-asymmetric filtering patterns.

I propose in chapter 6 a novel representation of questions based on a trivalent extension of inquisitive semantics. The general idea is that polar questions should receive a denotation that distinguishes positive and negative answers, in a parallel way to how declaratives are analysed. This allows us to apply trivalent connectives, as it has been proposed for declaratives, to questions, leading to an account of question coordination where presupposition projection and answerhood conditions are tightly linked. Importantly, the connectives are sensitive to the distinction between positive and negative answers, and therefore could not be defined purely in terms of the answer set.

Chapter 7 develops a second theory based on the same idea of making questions more similar to propositions, this time within a dynamic framework. I distinguish positive and negative answers by introducing bilateralism to the dynamic system, as Krahmer and Muskens (1995) propose as a solution to the problem of so-called bathroom sentences. The resulting theory is very high in the hierarchy shown in figure 1.1, and I show that it allows us to define the same sort of connectives that go beyond alternatives as in chapter 7, as well as allowing for various analytical possibilities of the sort offered by categorial theories.

References

Aloni, Maria and Robert van Rooy (2002). "The dynamics of question and focus". In: *Semantics and Linguistic Theory*. Vol. 12, pp. 20–39.

Alonso-Ovalle, Luis and Aron Hirsch (2018). "Keep *only* strong". In: *Semantics and Linguistic Theory 28*. Ed. by Sireemas Maspong et al., pp. 251–270.

Baroni, Marco (2021). "On the proper role of linguistically-oriented deep net analysis in linguistic theorizing". Ms. Facebook AI Research. Accessed June 2021 at lingbuzz/006031.

Bergen, Leon and Noah D. Goodman (2015). "The strategic use of noise in pragmatic reasoning". *Topics in cognitive science* 7.2, pp. 336–350.

Biezma, Maria and Kyle Rawlins (2012). "Responding to alternative and polar questions". *Linguistics and Philosophy* 35.5, pp. 361–406.

Bolinger, Dwight (1978). "Yes-no questions are not alternative questions". In: *Questions*. Ed. by Henry Hiż. Springer, pp. 87–105.

Breheny, Richard et al. (2018). "The symmetry problem: Current theories and prospects". *Natural Language Semantics* 26, pp. 85–110.

Buccola, Brian and Andreas Haida (2017). "Obligatory irrelevance and the computation of ignorance inferences". Ms. (July 31, 2017); lingbuzz/003600.

Buccola, Brian and Andreas Haida (2019). "Obligatory irrelevance and the computation of ignorance inferences". *Journal of Semantics* 36.4, pp. 583–616.

Büring, Daniel and Christine Gunlogson (2000). "Aren't positive and negative polar questions the same?" Ms., UCSC/UCLA.

Chemla, Emmanuel and Benjamin Spector (2011). "Experimental evidence for embedded scalar implicatures". *Journal of Semantics* 28.3, pp. 359–400.

Chierchia, Gennaro, Danny Fox, and Benjamin Spector (2011). "The grammatical view of scalar implicatures and the relationship between semantics and pragmatics". In: *Semantics: an international handbook of natural language meaning*. Ed. by Paul Portner, Claudia Maienborn, and Klaus von Heusinger. Vol. 3. De Gruyter Mouton, pp. 2297–2332.

Ciardelli, Ivano, Jeroen Groenendijk, and Floris Roelofsen (2013). "Inquisitive Semantics. A New Notion of Meaning". *Language and Linguistics Compass* 7.9, pp. 459–476.

Cummins, Chris, Uli Sauerland, and Stephanie Solt (2012). "Granularity and scalar implicature in numerical expressions". *Linguistics and Philosophy* 35.2, pp. 135–169.

Dotlačil, Jakub and Floris Roelofsen (2019). "Dynamic inquisitive semantics: Anaphora and questions". In: *Proceedings of Sinn und Bedeutung 23*. Ed. by M.Teresa Espinal et al. University of Konstanz, pp. 365–382.

von Fintel, Kai and Sabine Iatridou (2007). "Anatomy of a modal construction". *Linguistic Inquiry* 38.3, pp. 445–483.

Fox, Danny (2018). "Partition by exhaustification: Comments on Dayal 1996". In: *Proceedings of Sinn und Bedeutung*. Vol. 22. 1, pp. 403–434.

Fox, Danny and Martin Hackl (2006). "The universal density of measurement". *Linguistics and Philosophy* 29.5, pp. 537–586.

Fox, Danny and Roni Katzir (2011). "On the characterization of alternatives". *Natural Language Semantics* 19.1, pp. 87–107.

Fox, Danny and Roni Katzir (2020). "Notes on iterated rationality models of scalar implicatures". Ms. Massachussetts Institute of Technology and Tel Aviv University. Accessed June 2021 at lingbuzz/005519.

Fox, Danny and Benjamin Spector (2018). "Economy and embedded exhaustification". *Natural Language Semantics* 26.1, pp. 1–50.

Franke, Michael (2011). “Quantity implicatures, exhaustive interpretation, and rational conversation”. *Semantics and Pragmatics* 4.1, pp. 1–82.
Gärdenfors, Peter (2004). *Conceptual spaces. The geometry of thought*. MIT press.
Geurts, Bart and Frans van der Slik (2005). “Monotonicity and processing load”. *Journal of Semantics* 22.1, pp. 97–117.
Goodman, Noah D. and Andreas Stuhlmüller (2013). “Knowledge and implicature. Modeling language understanding as social cognition”. *Topics in Cognitive Science* 5.1, pp. 173–184.
Grice, Paul (1991). *Studies in the way of words*. Harvard University Press.
Groenendijk, Jeroen and Martin Stokhof (1984). “Studies on the semantics of questions and the pragmatics of answers”. PhD thesis. University of Amsterdam.
Haida, Andreas (2007). “The Indefiniteness and Focusing of *Wh*-Words”. PhD thesis. Humboldt University, Berlin.
Hamblin, Charles L. (1958). “Questions”. *Australasian Journal of Philosophy* 36.3, pp. 159–168.
Hamblin, Charles L. (1976). “Questions in Montague English”. In: *Montague grammar*. Ed. by Barbara H. Partee. Academic Press, pp. 247–259.
Horn, Laurence R. (1973). “On the semantic properties of logical operators in English”. PhD thesis. University of California.
Kao, Justine T. et al. (2014). “Non-literal understanding of number words”. *Proceedings of the National Academy of Sciences* 111.33, pp. 12002–12007.
Karttunen, Lauri (1973). “Presuppositions of compound sentences”. *Linguistic Inquiry* 4.2, pp. 169–193.
Karttunen, Lauri (1977). “Syntax and semantics of questions”. *Linguistics and philosophy* 1.1, pp. 3–44.
Katzir, Roni and Raj Singh (2013). “Constraints on the lexicalization of logical operators”. *Linguistics and Philosophy* 36.1, pp. 1–29.
Klinedinst, Nathan (2004). “Only scalar *only*”. Handout.
Krahmer, Emiel and Reinhard Muskens (1995). “Negation and disjunction in discourse representation theory”. *Journal of Semantics* 12.4, pp. 357–376.
Krasikova, Sveta (2010). “Sufficiency inference in anankastic conditionals”. In: *Proceedings of SALT 20*. Ed. by Nan Li and David Lutz, pp. 91–108.
Krifka, Manfred (2001). “For a structured meaning account of questions and answers”. In: *Audiatur vox sapientia. A festschrift for Arnim von Stechow*. Ed. by Caroline Féry and Wolfgang Sternefeld. Akademie Verlag, pp. 287–319.
Lassiter, Daniel and Noah D. Goodman (Oct. 2017). “Adjectival vagueness in a Bayesian model of interpretation”. *Synthese* 194.10, pp. 3801–3836.
Li, Haoze (2019). “Deriving short answers from Hamblin sets: A case study of Mandarin *wh*-conditionals”. In: *Proceedings of SALT 29*. Ed. by Katherine Blake et al. Linguistic Society of America, pp. 138–159.
Magri, Giorgio (2009). “A Theory of individual-level predicates based on blind mandatory scalar implicatures”. *Natural Language Semantics* 17.3, pp. 245–297.
Magri, Giorgio (2017). “Blindness, Short-sightedness, and Hirschberg’s contextually ordered alternatives: a reply to Schlenker (2012)”. In: *Linguistic and psycholinguistic approaches on implicatures and presuppositions*. Ed. by Salvatore Pistoia-Reda and Filippo Domaneschi. London: Palgrave Macmillan, pp. 9–54.
Penka, Doris (2012). “Split scope of negative indefinites”. *Language and Linguistics Compass* 6.8, pp. 517–532.

Qing, Ciyang and Michael Franke (2014). “Gradable adjectives, vagueness, and optimal language use. A speaker-oriented model”. In: *Semantics and Linguistic Theory (SALT) 24*, pp. 23–41.

Roberts, Craige (1996). “Information structure in discourse: towards an integrated formal theory of pragmatics”. *OSU working papers in linguistics* 49, pp. 91–136.

Romero, Maribel and Chung-Hye Han (2004). “On negative yes/no questions”. *Linguistics and Philosophy* 27.5, pp. 609–658.

Rooth, Mats (1992). “A theory of focus interpretation”. *Natural Language Semantics* 1.1, pp. 75–116.

Sauerland, Uli (2004). “Scalar implicatures in complex sentences”. *Linguistics and philosophy* 27.3, pp. 367–391.

Schlenker, Philippe (2008). “Be articulate: A pragmatic theory of presupposition projection”. *Theoretical Linguistics* 34.3, pp. 157–212.

Spector, Benjamin (2008). “An unnoticed reading for *wh*-questions: elided answers and weak islands”. *Linguistic Inquiry* 39.4, pp. 677–686.

Spector, Benjamin (2016). “Comparing exhaustivity operators”. *Semantics and Pragmatics* 9.11, pp. 1–33.

Sudo, Yasutada (2013). “Biased polar questions in English and Japanese”. In: *Beyond Expressives: Explorations in Use-Conditional Meaning*. Ed. by Daniel Gutzmann and Hans-Martin Gärtner. Brill, pp. 275–295.

Villalta, Elisabeth (2008). “Mood and gradability: an investigation of the subjunctive mood in Spanish”. *Linguistics and Philosophy* 31.467.

Xiang, Yimei (2021). “A hybrid categorial approach to question composition”. *Linguistics and Philosophy* 44, pp. 587–647.

Chapter 2

On the alternatives and implicatures of comparative modified numerals

This chapter was originally published as:

> Émile Enguehard (2018). "Comparative modified numerals revisited: scalar implicatures, granularity and blindness to context". In: *Proceedings of the 28th Semantics and Linguistic Theory Conference (SALT 28)*. Ed. by Sireemas Maspong et al., pp. 21–39.

Abstract

I present a theory of the pragmatics of comparative modified numerals such as "more than 5". The proposal is based on the assumption that they trigger alternatives derived from discrete "granularity scales". This is sufficient to explain the basic pattern of scalar implicatures from comparative numerals; I then show how extra assumptions of blind exhaustification and QUD uncertainty refine the prediction in that we can explain how and when comparative numerals trigger irrelevance inferences.

Acknowledgements

I thank Benjamin Spector and the audience at SALT 28 for discussion and comments.

2.1 Introduction

Comparative modified numerals, such as "more than five", are known to not trigger the scalar implicatures that bare numerals do. For instance, (1) is generally read to mean that John doesn't own more than three cars, which is often explained in the literature by a scalar implicature due to the presence of an alternative with "four":

(1) John owns three cars.

Instead, unembedded comparative numerals give rise to so-called ignorance inferences (the speaker doesn't know the number) or to irrelevance inferences (the question under discussion makes the precise number irrelevant). This is seen in (2):

(2) John owns more than three cars.
$\not\rightsquigarrow$ John doesn't own more than four cars.

Cummins, Sauerland, and Solt (2012) note that comparative modified numerals may in fact receive enriched readings when they combine with "round" numerals, as in (3).

(3) There are more than 90 people in this room.
$\rightsquigarrow$ There aren't more than 100.

We can extend this observation to non-round numbers. Specifically, enriched readings of non-round comparative numerals are sometimes available when they refer to continuous quantities, as in (4).

(4) a. John walked more than 7 kilometers to get home.
$\rightsquigarrow$ He didn't walk 8 kilometers.
b. John has been working there for more than 22 years.
$\rightsquigarrow$ He hasn't been there for 23 years.

2.2 Comparative numerals in the literature

The most straightforward approach towards comparative numerals, exemplified by Mayr 2013, takes them to evoke alternatives such as "exactly n" or "more than n" for integer n. Those alternatives are used in an exhaustification mechanism that evaluates them in context, resulting in the prediction that (2) triggers the *ignorance inference* that the speaker thinks it's possible that John read only four books, or that he read more.

Another approach is that of Fox and Hackl (2006); they introduce the *Universal Density of Measurement* (UDM) hypothesis, according to which the alternatives of modified numerals are based on all real numbers. Together with the assumption that the mechanism giving rise to scalar implicatures is blind to contextual information, they predict that comparative numerals cannot receive enriched meanings, not even ignorance implicatures.

Neither approach provides an immediate explanation for the full pattern we outlined above. The theory I am going to present is based on discrete scales; in section 2.6.3, we'll see how equivalent predictions could be derived from dense scales.

2.3 First proposal

2.3.1 Formal exposition

To capture the difference between (2) and (3), we need to make reference to a notion of *granularity* in our definition of alternatives. This is provided by Cummins, Sauerland, and Solt's (2012) *granularity scales*. A granularity scale is the scale of all numerals of a certain "roundness" such as:

a. $\langle 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, \ldots\rangle$

b. $\langle 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, \ldots\rangle$

c. $\langle 0, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, \ldots\rangle$
etc.

We assume that every modified numeral is interpreted within a granularity scale, and this determines its formal alternatives. Formally, if n is interpreted with the granularity scale S:

$$\textsc{alt}(\text{“more than } n\text{”}) = \{\text{“}m\text{”}, \text{“more than } m\text{”}, \text{“exactly } m\text{”} \mid m \in S\} \qquad (A)$$

where the semantics of "n" are assumed to be "n or more". The granularity scale chosen is the coarsest one available for the numeral (the units for "23", tens for "20", etc.).[1] Complete utterances also have formal alternatives which are obtained through replacing any of their constituents by their own formal alternatives.

We assume that any utterance answers a tacit "Question Under Discussion" (QUD), and for a given QUD Q there is a fixed set of "relevant" propositions R_Q. Four our purposes, we will consider two kinds of question:

1. "How many?" questions like "How many cars does John own?". We will call these *total questions*, and assume they make all alternatives we discuss relevant.

2. *Partial questions* like "Does John own more than three cars?". We will assume these questions make all other answers than (2) and its negation irrelevant.

Finally, we adopt the following fairly standard treatment of implicatures, except that we treat them as obligatory:

1. If ϕ answers Q, for all $\psi \in \textsc{alt}(\phi) \cap R_Q$ such that ϕ doesn't entail ψ, the hearer understands that the speaker doesn't believe that ψ is true (primary implicature). The set of such propositions is $\textsc{alt}'(\phi; Q)$.

2. If $\psi \in \textsc{alt}'(\phi; Q)$ is innocently excludable in the sense of Fox 2007,[2] the hearer may infer that the speaker believes that ψ is false (secondary implicature).

[1] Possibly the set of available granularity scales is contextually or culturally dependent, though there is no clear evidence for or against it. Possibly too, context may sometimes select a more precise granularity scale; see section 2.6. Finally, in some cases the "coarsest scale available" might be poorly defined, as it seems that non-linear scales like $\langle 0, 1, 10, 100, \ldots\rangle$ are available (Cummins, Sauerland, and Solt 2012).

[2] ψ is innocently excludable within $\textsc{alt}'(\phi; Q)$ if and only if every maximal co-deniable subset of $\textsc{alt}'(\phi; Q)$ contains ψ. A subset is co-deniable if and only if the conjunction of the negations of all its elements isn't a contradiction. Intuitively, this means that ψ is innocently excludable if it can be taken to be false together with any other alternatives without resulting in a contradiction.

In a given context, there may be several salient questions to which ϕ is a relevant answer. Upon hearing ϕ, the hearer then gets a disjunctive inference: for at least one of the salient questions, the primary and secondary implicatures as computed above hold, and the alternatives not in R_Q are irrelevant to the speaker's intention. Certain choices of Q may be eliminated if they result in pragmatically odd meanings.

2.3.2 Applying the mechanism

We will now look at how the theory above applies to (2) and (3). Per (A), the alternatives to (3) are those given in (5). Those that are neither weaker than nor incompatible with (3) are all those that involve numerals above 100, as in (6). Among those there is a single weakest element (7), involving a bare numeral.

(5) There are (more than / exactly / ∅) (.../80/90/100/...) people in this room.

(6) There are (more than / exactly / ∅) (100/110/120/...) people in this room.

(7) There are 100 people in this room.

If the most salient QUD is the total one, which makes all alternatives relevant, step 1 and step 2 above result in the listener's inference that the alternatives in (6) are false, as it can be verified that they are all innocently excludable. This is equivalent to just inferring that the weakest one, (7), is false. Thus we get the enriched meaning: there are more than 90 people, but fewer than 100.

If the most salient QUD is the partial QUD "are there more than 90 people here?", all alternatives other than (3) are irrelevant. We predict (3) to not receive an enrichment.

If there is uncertainty about the QUD, we predict a disjunctive inference: either (7) is false, or all the propositions in (6) are irrelevant.

Those predictions are consistent with the fact that (3) is acceptable when there are clearly much more than 90 people only if the partial QUD is salient, as in (8a), while it is quite degraded under an explicit total QUD as in (8b).

(8) Context: it is visible that there are around 200-300 people.
  a. A: We've got 90 chairs.
     B: There are more than 90 people here.
  b. A: How many people are there?
     #B: There are more than 90 people here.

The treatment of (3) extends to any usage of round numerals or any description of dense quantities. For instance, the case of (4a) would follow the same pattern, only using the unit scale instead of the tens. The alternatives of (4a) are listed in (9), with the ones that are non-weaker and co-tenable with (4a) in (10). This set has exactly the same structure as (6); there is a weakest element (11), counterpart to (7).

(9) John walked (more than / exactly / ∅) (.../6/7/8/...) kilometres to get home.

(10) John walked (more than / exactly / ∅) (8/9/10/...) kilometres to get home.

(11) John walked 8 kilometres to get home.

Without repeating the detail of it, we predict for (4a) a disjunctive inference completely parallel to the one we predict for (3): either all the alternatives in (10) are irrelevant, or (11) is false.

The other case we want to account for is that of non-round numerals being used to describe discrete quantities, as in (2). The alternatives to (2) given by (A) are in direct correspondence to those of the previous example (3); they are listed in (12). What changes is that we no longer have a single minimal non-weaker alternative: the bare numeral counterpart to (7), given in (13), is not non-weaker than the initial proposition (2): if John owns more than three cars, he owns four cars. Instead, among those alternatives that are non-weaker than and not incompatible with (2), there are two minimal elements, (14a) and (14b). They are minimal in the sense that there are no elements in the set that are weaker than either.

(12) John owns (exactly / more than / ∅) (1 / 2 / 3 / 4 / ...) car(s).

(13) John owns four cars.

(14) a. John owns exactly four cars.
b. John owns more than four cars.

(14a) and (14b) are not innocently excludable: given (2), negating either one entails that the other is true. This is the well-known pattern of symmetric alternatives. Hence, no secondary implicature should arise from them: if the salient QUD is the total question, which makes all alternatives relevant, we predict an *ignorance inference*. The speaker doesn't know whether John has exactly four or more than four cars. This is also what Mayr (2013) predicts, and for the same reasons.

As before, if the salient QUD is the partial question, we predict a non-enriched reading. If both QUD are salient, we predict a disjunctive inference: either the speaker is ignorant about (14a) and (14b), or they are irrelevant.

To conclude this section, we can note that the key property that distinguishes (3) and (4a) from (2) is whether there are possible worlds where the quantity being discussed is both above the stated bound and below the next item on the granularity scale. For instance, it is possible to walk more than 7 and less than 8 kilometres, or for there to be more than 90 and less than 100 people, but it is not possible to own more than 3 and less than 4 cars. This is why we predict a secondary scalar implicature for (4a) and (3), but not for (2). The fact that (3) is about a dense quantity and (4a) is about a discrete one has no consequence in our theory.

2.4 Refining the proposal: structural entailment and obligatory irrelevance inferences

In the discussion so far, we have adopted the contextual definition of entailment: ϕ (contextually) entails ψ if all contextually likely worlds where ϕ is true are also worlds where ψ is true. However, certain theories, including the UDM treatment of modified numerals by Fox and Hackl (2006), make use of a stronger notion of entailment, *structural* (or *logical*) entailment. ϕ structurally entails ψ if it is true at a smaller set of worlds, i.e. if all worlds that make ϕ true also make ψ true, independently of the current context. In the rest of this section, I will argue for structural entailment by showing how it modifies the predictions of the proposal, and why such changes are for the better.

Recall that propositions containing "more than n" and involving round numerals or dense quantities have a single minimally stronger alternative, containing "$n+1$ (or more)." When non-round numerals are used to describe discrete quantities, the corresponding alternative is equivalent to the original proposition in the sense of contextual entailment.

For instance, (15) (repeated from (2)) has the alternatives in (16) (ignoring strictly weaker and incompatible ones). The weakest among them is (17), which is contextually equivalent to (15). Because it is equivalent to the original proposition (and therefore weaker in the general sense), it isn't considered by the procedure described in section 2.3.1.

(15) John owns more than three cars.

(16) John owns (more than / exactly / ∅) (4/5/6...) cars.

(17) John owns four cars.

With a structural definition of entailment, it is no longer true that (17) is equivalent to (15), as long as we take the fact that cars come in integer quantities to be peculiar to the context.[3]

1. Our prediction under the partial QUD isn't affected by this fact, since all alternatives are irrelevant anyway.

2. Under the total QUD, all alternatives are relevant. They are also all innocently excludable, and denying them is equivalent to denying the single minimal one, namely (17). Thus, we predict the following enriched meaning: John owns more than three cars, but fewer than four. This is a contextual contradiction: the sentence should be infelicitous.[4]

3. When there is uncertainty about the QUD, we ought to predict again a disjunctive inference: either John owns fewer than 4 cars, or the number of cars he owns is irrelevant. Because the first disjunct is impossible, the listener has to accept the second one. Hence, what we actually predict is an obligatory *irrelevance inference*: the listener concludes that the speaker is interested in the partial QUD, even if it weren't very salient.

Compared to the contextual definition of entailment, we no longer predict the possibility of ignorance inferences. This is at odds with some of the literature, such as Mayr 2013, where such ignorance inferences are taken to be fact, and the theory is designed to derive them. Buccola and Haida (2017) are among some that argue that they don't exist. Indeed, in dialogues where the total QUD is made very explicit, such as (18), while "at least" numerals are felicitous even with non-round numbers, "more than" numerals are somewhat degraded. Given that "at least" numerals are well-known to trigger ignorance inferences (Schwarz 2016), this suggests that "more than" numerals are different. If, as we predict, they trigger obligatory selection of the partial QUD, their infelicity as an answer to the total one is readily explained.

(18) A: How many dogs does Ann own? Three? Seven? Twelve? I have no idea.

  a. B: At least 5 / 10 / 13.

[3]There are plausible worlds where John owns a fraction of a car, say if he shares it with a relative. I do not think we should rely on this fact, though: first, it doesn't intuitively spring to mind; second, it may not be true in a given time and place, say for legal reasons, and yet this shouldn't affect the linguistic facts; third, all what we say about (2) also applies to "Hexagons have more than 5 sides" as far as I can tell; yet, it is hard to imagine what a shape having a fraction of a side would correspond to.

[4]Note that even the primary implicature is already a contextual contradiction: John owns more than three cars, but the speaker isn't sure he owns four. Our prediction doesn't depend on secondary implicatures being obligatory.

b. B: More than ??5 / 10 / ??13. (adapted from Buccola and Haida 2017)

Another case where comparative numerals are degraded is when other partial QUDs, involving other thresholds, are very salient in the context. For instance, in the context of (19), 18 (or 21) is a much more relevant threshold than 33, making the accommodation of the partial QUD centred on 33 difficult. The superlative numeral is felicitous as it is a valid answer to the total QUD, which is, presumably, always relevant.

(19) (We're in a bar.) (Benjamin Spector (p.c.))
Of course she can drink, she's (at least / #more than) 33.

Given such examples, the choice of structural entailments, without changing our architecture, seems to provide better empirical adequacy.

To summarize the pattern that our theory predicts under both definitions of entailment:

1. With both definitions, we predict two readings when the quantity in question could be measured on a more precise scale as in (3), either because round numerals are used or because the quantity is dense. There is first an enriched reading with a scalar implicature, associated to a total QUD ("how many?"), and second, a non-enriched reading associated to a partial QUD. This seems empirically valid.

2. With the contextual definition, when the quantity cannot be measured any more precisely, because non-round numerals are used about discrete quantities, we predict no secondary implicature, plus an ignorance inference only when the total QUD is salient.

3. In the structural variant, when the quantity cannot be measured any more precisely, we predict no secondary implicature, and an obligatory irrelevance inference: the listener has to accommodate the partial QUD regardless of context. Evidence points in favour of this prediction over the one above.[5]

2.5 Embedded environments

2.5.1 Universal quantification

In this section, we will explore the predictions of our proposal when universal quantifiers are involved, and find further evidence teasing apart contextual and structural entailment.

To begin with, we will look at the case of discrete quantities and non-round numerals as in (20). Its alternatives (omitting strictly weaker and incompatible ones) are listed in (21). This set includes a logically minimal element, (22). The rest of the set has two incomparable minimal elements, (23a) and (23b).

(20) Every student of mine made more than 7 mistakes.

(21) Every student of mine made (more than / exactly / ∅) (8/9/10/...) mistakes.

(22) Every student of mine made 8 mistakes.

(23) a. Every student of mine made more than 8 mistakes.
b. Every student of mine made exactly 8 mistakes.

[5]Note that while the structural theory doesn't allow for an ignorance inference, the reading it predicts is *compatible* with speaker ignorance.

Assume a QUD that makes all alternatives relevant is salient. Under the contextual definition of entailment, (22) is equivalent to (20), and therefore weaker in a general sense. We only derive implicatures from non-weaker alternatives, so (22) is ignored. All other alternatives are innocently excludable, and we predict them to trigger secondary implicatures. Because denying them all is equivalent to denying the two minimal ones, (23a) and (23b), the listener's final inference upon hearing (20) is that not all students made exactly 8 mistakes, nor did they all make more than 8 mistakes. In other words, the enriched meaning of (20) is that some students did 8 mistakes exactly, and some did more.

Under the structural definition of entailment, (22) is non-weaker than (20). Again, all alternatives in (21) are innocently excludable, and denying them all is tantamount to denying (22). Hence, we predict an inference that some students did fewer than 8 mistakes. This is a contextual contradiction, and (20) ought to be infelicitous. However, if a partial QUD is at least somewhat salient in the context, we predict it to be obligatorily selected (irrelevance inference).

Thus, there is no difference with the unembedded case as far as the structural definition of entailment is concerned. For the contextual definition, we get a dual secondary implicature instead of an ignorance inference.

Again, we can try to find out whether this dual implicature is possible by comparing to "at least $n+1$", for which this dual implicature is well-known to be there. If they triggered the same inferences, we would expect B's utterance in (24b) to be an equally valid continuation as B's utterance in (24a). Though judgements aren't clear, there appears to be a contrast. Such data point towards the structural definition of entailment as the better one within our proposal.

(24) Context: you need to make fewer than 10 mistakes to pass.

a. A: Every student of mine made at least 7 mistakes.
B: Well, at least some of them passed.

b. A: Every student of mine made more than 7 mistakes.
?B: Well, at least some of them passed.

In the case of dense quantities and round numerals, the definition of entailment makes no difference. (25) has the alternatives in (26), amongst which there is a single minimal element (27).

(25) There were more than 80 people at every class.

(26) There were (more than / exactly / ∅) (90/100/...) people at every class.

(27) There were 90 people at every class.

If all alternatives in (26) are relevant, since they are all innocently excludable, we may derive secondary implicatures from all of them. The inference this results in is the negation of (27): hence the enriched meaning we predict is that at certain classes there were fewer than 90 people.[6]

[6] The stronger inference that there were fewer than 90 people at every class is attractive, but it is easy to cancel, thus it cannot derive from the obligatory mechanism we're studying here:

(i) How many people came to the class?

a. There were more than 80 people at every class, and once they were 127!

b. #There were more than 80 people at every class, in fact the lowest number was 127.

2.5.2 Negation

We now describe the predictions of the mechanism when comparative numerals are embedded under negation.

We first have to specify what the alternatives of a sentence involving "not more than n." We will assume that if n is interpreted within the granularity scale S, then this phrase has the alternatives in (A').

$$\textsc{alt}(\text{“not more than } n\text{”}) = \{\text{“fewer than } m\text{”}, \text{“not more than } m\text{”}, \text{“exactly } m\text{”} \;/\; m \in S\} \quad (A')$$

The first two series of elements are negated forms of alternatives of "more than n", assuming that negating a bare numeral produces "fewer"; the third one is an alternative of "more than n".

We need to assume this is the alternative set to not make undesirable predictions: if "more than m" and "m (or more)" are included, as we would expect if deleting negation was a possible way of deriving alternatives, we run into the symmetry problem: there are no innocently excludable alternatives at all, and we predict ignorance inferences across the board; we will see how this doesn't match the data. If "exactly m" is not included, we predict a systematic enrichment of "not more than n" into meaning "exactly m"; we will discuss below whether this is really wrong.[7]

Let's first look at an example using a round numeral, (28). Its not strictly weaker nor incompatible alternatives, as derived from (A'), are listed in (29). Among them there are two minimal elements, (30a) and (30b). Among the rest, there is a single minimal element, (30c).

(28) There aren't more than 80 people here.

(29) a. There aren't more than .../60/70 people here.
b. There are fewer than .../60/70/80 people here.
c. There are exactly .../60/70/80 people here.

(30) a. There are fewer than 80 people here.
b. There are exactly 80 people here.
c. There aren't more than 70 people here.

As usual, under a partial QUD, all those alternatives are irrelevant and we predict no enrichment.

Given (28), negating (30a) entails that (30b) is true and vice-versa. This means that they are not innocently excludable. Hence, under a total QUD, we predict the listener to draw an ignorance inference about (30a) and (30b). However, (30c) and all stronger alternatives are innocently excludable and may trigger secondary implicatures. The full enrichment that we predict for (28) is therefore that there are between 70 and 80 people, and the speaker considers it possible that there are as many as 80, but isn't sure.

[7]A tentative justification for this necessary assumption is that deleting negation is allowed, but some alternatives are excluded because expressing lower bounds or upper bounds has an evaluative component. Specifically, sentences using "more than" are used when the quantity is perceived to be high, etc. This would explain why "more than" and "fewer than" do not compete.

We could avoid the assumption altogether if we assumed than the exhaustification procedure only makes reference to stronger alternatives (as opposed to non-weaker ones). Then "more than" and "fewer than" would not compete because there is no entailment relationship between them.

This is, as far as I can tell, empirically valid: B's answer in (31) is degraded because it violates the inference that there are more than 100 people,[8] and (32) is degraded when ignorance is difficult to accommodate.

(31) Context: there are clearly fewer than 50 people.
A: How many people are there here?
#B: There aren't more than 200.

(32) Context: I have 47.54€ on my account.
a. ? I just checked my balance, and I haven't got more than 50€.
b. I haven't checked my balance, but I haven't got more than 50€.

For the case of discrete quantities and non-round numerals, we will work off (33) as an example. Its alternatives (in (34)) have the same structure as those of (28): there are two minimal ones, (35a) and (35b), and the rest of them have a minimal element, (35c).

(33) John doesn't own more than 3 cars.

(34) a. John doesn't own more than 1/2 cars.
b. John owns fewer than 1/2/3 cars.
c. John owns exactly 1/2/3 cars.

(35) a. John owns fewer than 3 cars.
b. John owns exactly 3 cars.
c. John doesn't own more than 2 cars.

If entailment is defined contextually, (35c) is equivalent to (35a), and there are no innocently excludable alternatives. We predict (33) to be felicitous under the total QUD, and to trigger an ignorance inference.

If entailment is defined structurally, the computation starts out the same way as for (28). Under the total QUD, all alternatives are relevant. (35b) and (35a) are symmetric alternatives, and therefore not innocently excludable; we predict an ignorance inference about them. (35c) and all stronger alternatives are innocently excludable. The secondary implicature derived from (35c) leads to a contextual contradiction: John owns more than 2 cars and not more than 3, but the speaker isn't sure he owns exactly 3. On the other hand, if (35c) only results in a primary implicature, there is no contradiction. Thus, the prediction depends on whether we assume secondary implicatures are obligatory: if they are, we predict infelicity from the contextual contradiction; if they aren't, we predict felicity under the total QUD, together with an ignorance inference.

There is no evidence that "not more than *n*" is infelicitous under the total QUD; for instance, (36) is fine. Given that this is compatible with both definitions of entailment, this doesn't help us towards choosing between them.

(36) A: How many people will come? Five?
B: There won't be more than three of them.

[8]The primary implicature (B isn't sure there are fewer than 100 people) is in fact enough to explain why B's answer is degraded. Note that there is no ignorance inference about whether there are 100 people: B's answer is felicitous when there are clearly more than 100 people.

2.6 Issues

2.6.1 Further issues regarding negated comparative numerals

In this section, we describe various interesting issues that our proposal doesn't account for as well as might be hoped.

The first one is that the behaviour of negated comparative numerals is more complex than what we describe in section 2.5.2. To begin with, there are in fact two ways to negate comparative modified numerals in English, exemplified by (37) (repeated from (33)) and (38).

(37) John doesn't own more than 3 cars.

(38) John owns no more than 3 cars.

There is no reason why we would predict any difference between these two constructs. Yet, they are not completely equivalent. For instance, (37) is a fine answer to the partial QUD, in which case its meaning isn't enriched, as we predict. (38), though, is somewhat degraded as an answer to the explicit question "does John own more than 3 cars?," and it isn't clear what its meaning is. Similar facts are already described by Nouwen (2008), who notes that the "not" construct, where the negation has a wider surface scope and perhaps wider scope at LF too, is presumably more appropriate for "denial" — what we call answering the partial QUD here is probably a superset of denial.[9]

As an answer to the total QUD, "no more" can trigger the inferences that we predict: the listener draws an ignorance inference about whether the bound is reached for both discrete and dense quantities, as in (40) and (39), as well as a scalar implicature in the dense case, as in (40).

(39) John owns no more than 3 cars.
⇝ The speaker isn't sure John has as many as 3 cars, but they think it's possible. (possible reading)

(40) John drank no more than 2 litres of beer in his life.
⇝ John drank more than one litre, and the speaker isn't sure whether he reached two. (possible reading)

However, "no more" (and perhaps "not more" too) has an additional, stronger reading with no ignorance or irrelevance inference, which is the focus of Nouwen's (2008) analysis. This reading is one where the speaker is in fact fully answering the total QUD; for instance, (41) has a reading where the speaker believes that John has exactly 7 books, as shown by the validity of the continuation. This reading comes with a strong evaluative component: we understand that the speaker believes that the quantity in question is small.[10]

(41) John has no more than 7 books in his library.
(I counted them all.)

We would predict this reading if we didn't have the "exactly" alternative in (A'), which in turn would be the case if we assumed deleting negation in alternatives is not licit. Then,

[9] Another reason for this pattern could be that stressing the negation is a way to make the partial QUD more salient, and it is not possible in (38).

[10] In fact most examples in this paper carry a suggestion that the quantity is small when stating an upper bound and high when stating a lower bound, at least when they're interpreted under the total QUD.

(35a) would be the single minimal alternative to (33). A possible explanation for the difference between “no more” and “not more” would then be that the negation in “no more” is harder to delete than the sentential one; this merits further exploration. Alternatively, it could be that the “exactly” alternative is never present: then, all the examples we gave of readings where the speaker is ignorant would just be instances of primary implicatures not being strengthened, rather than obligatory ignorance inferences: it remains to be seen how those two predictions can be told apart[11] In any case, since the strong reading seems to be accessible only in a restricted set of contexts, those where the quantity is perceived to be very small, while the weaker reading where the speaker is ignorant is more generally accessible, I chose to select the alternatives set that leads to generating the weaker reading.

2.6.2 The case of necessity modals

In this section, we review a puzzle on the behaviour of comparative modified numerals embedded under deontic necessity modals. Because they have the same structure, we predict necessity modals to behave similarly to universal quantifiers like “every”, that is:

1. When using round numerals or talking about dense quantities, there is a secondary scalar implicature.
2. When using non-round numerals and talking about discrete quantities, there is an obligatory irrelevance inference.

This is indeed what we observe with epistemic modals. (42) triggers an irrelevance inference, and (43) a scalar implicature.

(42) (For all I know,) John must have more than 3 cars.
⇝ It is irrelevant how many he has.

(43) There must be more than 90 people here.
⇝ There may be fewer than 100.

However, some deontic examples appear to have readings that we do not predict. (44), for instance, should trigger an irrelevance inference: its minimal alternative is (45), and strengthening that would result in the nonsensical proposition that John may answer strictly less than 9 questions and pass. Instead, (44) is a felicitous description of a rule stating that John may pass by answering 9 questions or by answering more.

(44) John must answer more than 8 questions to pass.

(45) John must answer 9 questions to pass.

(46) a. John must answer 9 questions exactly to pass.
b. John must answer more than 9 questions to pass.

The actual reading of (44) is what our proposal would straightforwardly predict if we had kept to contextual entailment. Indeed, (45) is contextually equivalent to (44). This makes the

[11] Yet another explanation could be that the “exactly” alternative may *sometimes* not be available, though this is harder to explain; certainly it cannot be irrelevant while the “fewer than” one is. Possibly, when prior expectations about the quantity are that it should be much higher, the “exactly” alternative is overly informative as the “no more” one already locates us in a sufficiently small corner of the epistemic space; this can be formalised within existing game-theoretical models of pragmatics.

alternatives in (46) the minimal ones; as they are innocently excludable, the final prediction is that (44) triggers the inference that they do not hold.

This fact puts a dent in the otherwise strong empirical validity of the structural variant. It is possible to rescue it, but it requires some questionable assumptions on modality. The idea is that the problematic inference obtained by denying (45) is not as untenable as other problematic inferences we had before. If deontic modality is assumed to be partly blind to world knowledge, and in particular, if the fact that it is only possible to answer integer numbers of questions is not part of the deontic modal base, then denying (45) isn't contradictory even together with (44): (44) asserts that answering between 8 and 9 questions is compatible *with the rules*, regardless of whether it is feasible or not. It is in fact possible to express explicitly such nonsensical rules, as in (47):

(47) Context: There are 43 voters in a yes/no ballot.
Per the rules, the change will be rejected if there are at least 21.5 "no" votes.

Another problem arises when deontic modals embed comparative numerals describing dense quantities or using round numbers. For instance, (48) is predicted to trigger the inference that it isn't necessary to run 4km to get a prize. It actually has a stronger reading where running any distance above 3km is enough to get a prize.

(48) To get a prize, you must run more than 3km.

This strong reading is one we could generate most straightforwardly from dense scales, following Fox and Hackl's (2006) UDM hypothesis. Arguably, our proposal also generates it. Indeed, in deontic contexts, it is easy to select precise granularity scales even where coarse ones are available: this is shown by the fact that (49), the counterpart to (44) using the round number 10, doesn't necessarily get the reading that answering fewer than 20 questions can be sufficient: instead it can mean that answering 11 questions can be sufficient.

(49) John must answer more than 10 questions to pass.

While we have no explanation for this fact, we take it to imply that (48) can also be interpreted with a very precise granularity scale, such as $\langle 3, 3.001, 3.002, \ldots \rangle$. Then, we predict it to mean something like "running 3.001km is enough to get a prize", which is mathematically different from the reading we're looking for, but indistinguishable from it in practice.

2.6.3 Alternative theories

The theory we've discussed so far is based on stipulated discrete alternative scales, of the kind that are traditionally used in the literature. In this section, we attempt to sketch alternative theories that would at least capture the pattern in section 2.1, while being based on a single dense scale, as per Fox and Hackl's (2006) UDM hypothesis.

A first approach would be to assert that granularity depends on the QUD. Thus, instead of a single "total QUD", we would have a variety of them for different granularities. Upon hearing "more than 10" when talking about discrete quantities, listeners would infer that the QUD is either the partial one, or the total one with granularity ten, because the utterance wouldn't be a relevant answer to a coarser total QUD, and would be incoherent as an answer to a more precise one. The predictions would be the same as what we've reviewed here, with an additional one: using, say "more than 10" should trigger irrelevance inferences about alternatives involving 11, 12, etc., which seems like a good result. I haven't presented this theory as it wasn't my initial approach, and its formalization could perhaps present some challenges.

We can try to stay closer to Fox and Hackl's analysis by relying on closure conditions on the alternatives set to derive the lack of scalar implicature, in the spirit of Fox and Katzir 2011 (Danny Fox and Benjamin Spector, p.c.). Assume that alternatives derived from the entire dense scale are part of the basic alternative set, before relevance considerations come into play. Assume also the following closure condition: alternatives that are contextually equivalent to the utterance have to be relevant. When describing a dense quantity, "more than 7 kms" and "more than 7.5 kms" are not contextually equivalent: we predict it should be possible for the latter alternative to be pruned, keeping only the "8 (or more)" alternative and deriving a secondary scalar implicature from it. On the other hand, when talking about discrete quantities, "more than 3.5 cars" and "4 (or more) cars" are contextually equivalent to "more than 3 cars": therefore they cannot be pruned, and exhaustification cannot take place.

While this theory predicts the lack of scalar implicatures for (2) (repeated below as (50)) and their presence for (4a) (repeated as (51)), it doesn't derive the obligatory irrelevance inference, and it doesn't explain why round numerals (as in (3), repeated as (52)) behave like dense quantities. Indeed, recall that our proposal makes no difference between (51) and (52) because the crucial property for us is whether it is tenable to deny the next bare numeral alternative: John may walk more than 7 and fewer than 8kms, there may be more than 90 and fewer than 100 people, and John may not own more than 3 and fewer than 4 cars. In the alternative proposal we just sketched, the crucial property is whether some alternatives are contextually equivalent to the utterance: John walking more than 7.1kms isn't equivalent to John walking more than 7kms, but John owning more than 3.1 cars is equivalent to John owning more than 3 cars, and there being more than 90.1 people is equivalent to there being more than 90 people. For the same reason that it isn't licit to rule irrelevant whether John owns more than 3.1 cars, it shouldn't be licit to rule irrelevant whether there are more than 90.1 people. This groups (52) together with (50), which is wrong.

(50) John owns more than 3 cars.

(51) John walked more than 7 kilometres to get home.

(52) There are more than 90 people here.

In the hypothetical theories we just sketched, the possibility of interpreting within the full dense scale should help account for problematic cases involving deontic modals such as (48), though one would need to explain why epistemic cases like (43) behave differently. The case of (44) is problematic for all the theories we discussed.

2.7 Discussion and perspectives

As we've seen, the combination of a structural view of entailment and the introduction of granularity scales into theories of alternatives lets us provide fairly comprehensive empirical coverage of comparative numerals.[12] We can explain:

[12] I haven't given any example involving a "fewer than" numerals, but if their alternatives are those in (A'), the predictions we make for them are completely symmetric to those we make for "more than" numerals, which seems correct.

a. Why they don't trigger ignorance implicatures, in contrast with so-called superlative numerals.

b. Why and when they trigger irrelevance implicatures.

c. Why and when they trigger standard scalar implicatures (contrary to previous assumptions that they never did).

Straightforward theories based on discrete alternative scales (including our initial "contextual" proposal) fail to explain the first point, while in theories based on dense scales, as per Fox and Hackl's (2006) UDM hypothesis, it isn't straightforward to explain why scalar implicatures occur in cases involving discrete quantities. Our proposal doesn't directly bases its predictions on whether the quantity being discussed is dense.

One of the important ingredients of this proposal, or of its hypothetical alternative formulations, is mandatory blind exhaustification. Mandatory blind exhaustification has been proposed for other phenomena as well (cf. for instance Magri 2009), but it runs against the common idea that implicatures are cancellable. The architecture of our model, where different exhaustifications are possible depending on the QUD, may perhaps be used to explain this fact in more general account of implicatures, where we'll try to characterize cancellable and "pragmatic" inferences as deriving from QUD selection. It remains to be seen how far this can be pushed.

As far as modified numerals are concerned, the natural next step from here is to address the behaviour of superlative numerals, and first, assess their interaction with granularity. Besides, an important trait of modified numerals that has only been briefly touched on in this paper is the presuppositions or evaluative component on prior expectations and desires that they carry. This is particularly true for expressions like "only *n*" or, as we mentioned, "no more than *n*" and (even more so) "no less/fewer than *n*", in their "exactly" readings. The formalism we have been using here doesn't seem to be able to account for such effects: it would be desirable to find out how to solve this, and, to begin with, to fully examine the interaction between the sort of implicatures we're interested in and a-priori expectations.

References

Buccola, Brian and Andreas Haida (2017). "Obligatory irrelevance and the computation of ignorance inferences". Ms. (July 31, 2017); lingbuzz/003600.

Cummins, Chris, Uli Sauerland, and Stephanie Solt (2012). "Granularity and scalar implicature in numerical expressions". *Linguistics and Philosophy* 35.2, pp. 135–169.

Fox, Danny (2007). "Free Choice Disjunction and the Theory of Scalar Implicatures". In: *Presupposition and Implicature in Compositional Semantics*. Ed. by Uli Sauerland and Penka Stateva. Palgrave Macmillan, pp. 71–120.

Fox, Danny and Martin Hackl (2006). "The universal density of measurement". *Linguistics and Philosophy* 29.5, pp. 537–586.

Fox, Danny and Roni Katzir (2011). "On the characterization of alternatives". *Natural Language Semantics* 19.1, pp. 87–107.

Magri, Giorgio (2009). "A Theory of individual-level predicates based on blind mandatory scalar implicatures". *Natural Language Semantics* 17.3, pp. 245–297.

Mayr, Clemens (2013). "Implicatures of modified numerals". In: *From Grammar to Meaning: the Spontaneous Logicality of Language*. Ed. by Ivano Caponigro and Carlo Cecchetto. Cambridge University Press, pp. 139–171.

Nouwen, Rick (2008). “Upper-bounded no more: the exhaustive interpretation of non-strict comparison”. *Natural Language Semantics* 16.4, pp. 271–295.
Schwarz, Bernhard (Apr. 2016). “Consistency preservation in Quantity implicature: The case of *at least*”. *Semantics and Pragmatics* 9.1, pp. 1–47.

Chapter 3

Connectedness as a constraint on exhaustification

Written with Emmanuel Chemla (LSCP, ENS/EHESS/CNRS). This chapter was originally published as:

> Émile Enguehard and Emmanuel Chemla (2019). "Connectedness as a constraint on exhaustification". *Linguistics and Philosophy* 44.1, pp. 19–112.

Abstract

"Scalar implicatures" is a phrase used to refer to some inferences arising from the competition between alternatives: typically, "Mary read some of the books" ends up conveying that Mary did not read *all* books, because one could have said "Mary read all books". The so-called grammatical theory argues that these inferences obtain from the application of a covert operator EXH, which not only has the capability to negate alternative sentences, but also the capability to be embedded *within* sentences under other linguistic operators, i.e. EXH has the potential to add to the meaning of expressions (not necessarily full sentences), the negation of their alternatives. This view typically seeks support from the existence of readings that could not be explained without the extra-capability of EXH to occur in embedded positions. However, if some embedded positions seem to be accessible to EXH, not all conceivable positions that EXH could occupy yield sensible results. In short: the EXH approach is powerful, maybe too powerful.

Various approaches based on logical strength and monotonicity have been proposed to justify on principled grounds the limited distribution of EXH; these approaches are mostly based on a comparison between possible parses, and considerations of monotonicity (e.g., the Strongest Meaning Hypothesis). We propose a new constraint based instead on "connectedness", ruling out parses because of inherent problems their outcome may raise. Connectedness is a sister notion of monotonicity, which has been recruited to explain certain lexical restrictions on nouns, adjectives and more recently quantifiers; we propose here that connectedness could play a similar role at the level of propositional meanings.

Acknowledgements

We would like to thank Brian Buccola and Benjamin Spector for important discussions.

3.1 The distribution of exhaustification

3.1.1 The EXH operator

In many occasions, (1) is understood to imply that John didn't do all of his homework. This inference is not assumed to follow from the literal meaning of (1), which basically only says positive things about John and his homework. Instead, this inference is traditionally viewed as an enrichment, which can be explained if one assumes that (1) has (2) as a natural alternative, and that, in appropriate circumstances, alternatives are inferred to be false. Concretely then, the result is an *enriched* or *exhaustified* meaning for (1), which not only implies that the literal meaning is true, but also that its alternative is false. The inference that (2) is false from hearing (1) may be called a *scalar implicature*.

(1) John did some of his homework.

(2) John did all of his homework.

To formalize this, one may define an exhaustification function, EXH, that takes as input a sentence ϕ and its set of alternatives A (indicated as a subscript), and outputs an enriched meaning which adds to the literal meaning the relevant negations of alternatives: The meaning of $\text{EXH}_A[\phi]$ then is the meaning of ϕ conjoined with the negations of elements in A.[1] In the so-called Gricean view, scalar implicatures arise from pragmatic reasoning over speakers' intentions, and this function EXH may thus be seen as a way to formally package the result of such a reasoning (cf. for instance van Rooij and Schulz 2004 or Schulz and van Rooij 2006). In this view, this operation most naturally applies to whole sentences, and (1) gets its enriched meaning as the result of the following application of EXH:

(3) $\text{EXH}_{\{\text{John did all of his homework.}\}}$ [John did some of his homework.]

In contrast, Chierchia, Fox, and Spector (2011) develop a view of implicatures as derived from the grammar, in which such a function EXH would be a grammatical operator. As such, the EXH operator would be capable of taking a more diverse array of inputs. It would still be able to take whole sentences as inputs, and to explain the above facts *ceteris paribus*, but additionally it may occur in embedded positions, within full sentences. One of the initial motivations for this grammatical view was indeed the attempt to account for cases in which meaning is modified in place, that is *some* seems to acquire a *some but not all* meaning from within the scope of linguistic operators. Chierchia, Fox, and Spector (2011) summarize a wealth of evidence for such embedded occurrences of EXH. They argue for instance that (4a) has a reading where an embedded item like "most" is understood as "most but not all". This can be explained if EXH can appear above "most", as in the parse given in (4b). Without embedded exhaustification, (4a) would be incoherent: professors who fail all of their students are said to both be fired and receive a new assignment.

(4) a. Every professor who fails most of the students will be assigned a new class, and every professor who fails all of the students will be fired.
(adapted from Chierchia, Fox, and Spector 2011)
b. Alleged parse: Every professor who fails EXH [most of the students] will be assigned a new class, and every professor who fails all of the students will be fired.

[1] For simplicity, we will consider that EXH negates all alternatives a such that $\phi \wedge \neg a$ is not a contradiction; this is essentially the lexical entry for "only" in Krifka (1993). See Fox (2007) and Spector (2016) among others for necessary refinements.

Other cases seem to call for more flexibility than just global applications of EXH, typical examples include cases in which a scalar item like “some” appears in the scope of quantifiers. This is illustrated in (5) and (6) below. There it seems that the output predicted by a global application of EXH is accessible, but the output predicted by the local application of EXH within the scope of the quantifier also seems to be accessible (see Chemla and Spector 2011 and Potts et al. 2016 for discussions about the exact relevance of these examples, as well as quantitative data about the availability of the alleged readings):

(5) Every employee saw some of the criminals.
  a. Possible parse: Every employee saw EXH [some of the criminals.]
     Resulting meaning: Every employee saw some of the criminals, but none saw all the criminals.
  b. Possible parse: EXH [Every employee saw some of the criminals.]
     Resulting meaning: Every employee saw some of the criminals, but not all of them saw all the criminals.

(6) Exactly 3 employees saw some of the criminals.
  a. Possible parse: Exactly 3 employees saw EXH [some of the criminals.]
     Resulting meaning: Exactly 3 employees saw some but not all of the criminals.
  b. Possible parse: EXH [Exactly 3 employees saw some of the criminals.]
     Resulting meaning: Exactly 3 employees saw some of the criminals, less than that saw all the criminals.

Another class of examples, also described by Chierchia, Fox, and Spector (2011), is that of so-called Hurford disjunctions. Here the argument is of a different nature: if we do not postulate the occurrence of an embedded EXH, then the sentence may violate independent linguistic principles. For instance, in (7), without EXH, the second disjunct is redundant: “all” is strictly stronger than “some,” and if A^+ is stronger than A, then A or A^+ is equivalent to A. With EXH however, the first disjunct of the form EXH[some] has a meaning akin to ‘some but not all’, which makes the two disjuncts logically independent. If A^+ and A' are logically independent, then $A' or A^+$ isn't equivalent to either A^+ or A', and there is no redundancy. Thus, postulating the possibility of embedding EXH in (7) can help explain why it does not violate a general ban against redundancy.[2]

(7) a. John did some or all of the homework.
  b. Possible parse: John did EXH [some] or all of the homework.

3.1.2 Empirical restriction on the distribution of EXH

At this point, the grammatical approach is formally more expressive than the neo-Gricean view, in that it predicts strictly more possible meanings. We have also shown cases in which its extra expressive power was welcome (though see Chemla 2009 for a case where it may miss the mark). We will now review cases in which this additional expressivity seems to be *too* powerful, in that it may allow for parses that produce unattested meanings.

Most famously, there is no reason at this point in the grammatical view why EXH could not appear under negation. But it is well-known that the readings this would lead to are at best highly dispreferred (see (15) however for a qualification). For instance, if sentence (8a)

[2] The analysis sketched here is not without challenges: see Katzir and Singh (2013b) for a more thorough discussion of Hurford disjunctions, and Mayr and Romoli (2016) for redundancy constraints in general.

could have a parse as in (8b), then it should have the meaning paraphrased in (8c), corresponding to 'some' being interpreted as 'some but not all'. Under negation, we obtain a meaning akin to *all or none*, which does not seem to be readily available.

(8) a. It's not true that John did some of his homework.
 b. Potential parse: *It's not true that John did EXH [some of his homework].
 c. Unattested meaning: John did all or none of his homework.[3]

Another classic example is that of three-way disjunctions, as in (9a). Assume that the sentence has the structure *A* or (*B* or *C*), and that "or" has the alternative "and". Then we should be able to obtain a meaning for it as in (9c), if we allow a parse as in (9b). This reading, as before, is unattested, suggesting that the corresponding types of embedded occurrences of EXH are banned.

(9) a. John will meet Ann, Bill or Carol.
 b. Potential parse: *EXH(*A* or EXH(*B* or *C*))
 c. Unattested meaning: John will meet exactly one or exactly three of those people.

We may also consider (10), a variation of (9). If parsed with an embedded EXH as in (10b), it ought to mean that John will either visit his mother, or do some but not all of his homework. Such a reading would make (10a) false when John doesn't visit his mother, and does all of his homework, but true if he does all of his homework and visits his mother. This reading, paraphrased in (10c), doesn't seem available.

(10) a. This afternoon, John will visit his mother or do some of his homework.
 b. Potential parse: *This afternoon, John will visit his mother or do EXH [some of his homework.]
 c. Unattested meaning: If John doesn't visit his mother, he won't do all of his homework.

Embedding of the EXH operator under existential quantifiers also seems to be difficult. In (11a), having an embedded EXH, as in (11b) again essentially turns "some" into "some but not all", resulting in the meaning paraphrased in (11c). Such an interpretation is true in a scenario where some employees saw just some criminals, and some other employees saw all of them, but it is false if some employees saw all criminals and the rest saw none. Such a reading seems quite marked, if it exists at all.

(11) a. There are employees who saw some of the criminals.
 b. Potential parse: *There are employees who saw EXH [some of the criminals.]
 c. Unattested meaning: There are employees who saw some but not all of the criminals.

[3] As an anonymous reviewer points out, the embedded EXH reading appears to become accessible when the scalar item bears prosodic prominence and a continuation follows, as in (i).

(i) It's not true that John did SOME of the homework. He did ALL of it.

Similar facts hold for many of our examples. Such cases have been called "intrusive implicatures" (e.g. Fox (2007)). We are chiefly concerned here with the distribution of *covert* exhaustification, and consider that intrusive implicatures, as they require marked prosody and specific continuations to obtain, fall outside the scope of what we seek to explain. In this respect we follow the earlier theories that we're going to discuss, as they are designed to ban (8b) in spite of the possibility of (i). The possibility of a link between our proposal and *overt* strategies for exhaustification is discussed briefly in Footnote 13.

3.1.3 Possible constraints on the distribution of EXH: Strongest Meaning Hypothesis (SMH) and Economy Constraint (EC)

In the defense of the grammatical view, one thus needs to preserve its extra-expressiveness that helps account for attested cases of embedded exhaustification (Section 3.1.1), but also avoid over-generalization and limit its power through constraints that explain the limited distribution of EXH (Section 3.1.2). These constraints should thus ban embedding of EXH under negation as in (8), within disjunctions as in (9) or (10), or under existential quantifiers as in (11). One candidate for such a constraint is the Strongest Meaning Hypothesis or SMH (Dalrymple et al. 1998). It has been stated and studied in the context of embedded EXH by Sauerland (2012) in the following form:

(12) **Strongest Meaning Hypothesis (SMH)**
A parse is dispreferred if there is another parse that differs only in terms of placement of EXH, and that leads to a stronger meaning.

The SMH can explain several of the above examples. Consider the example with negation in (8) first. Generally, EXH(A) is stronger than A, hence under negation ¬A is stronger than ¬EXH(A). As a result, the SMH systematically blocks the application of EXH under negation, and therefore bans the parse in (8b). Similar arguments hold for the cases with disjunctions, (9) or (10). Focussing on (9), the unavailable parse is of the form EXH(A or EXH(B or C)). As explained, this parse would yield the interpretation 'exactly one or exactly 3 of A, B and C hold'. This parse is blocked by the SMH because dropping the embedded EXH, as in EXH(A or (B or C)), yields the strictly stronger reading: 'exactly one of A, B and C holds'.[4] The case with existential quantification (11) would work similarly. In short, the parse with a matrix-level EXH results in the meaning that no employee saw all criminals; this is stronger than (11c), so the SMH predicts (11b) to be inaccessible.

The SMH is thus able to block some of the parses generated by the possibility to embed EXH, but yielding unattested readings. Does it annihilate the power of EXH completely however, in banning desirable parses too? Here the results are more complicated. Consider (6) first, with a scalar item embedded under the non-monotonic quantifier 'Exactly 3'. The two parses, with EXH embedded or not embedded, were found to yield available readings, and indeed the readings they produce (6a) and (6b) are logically independent and therefore survive the SMH. However, in other cases, the SMH makes unfortunate predictions and bans too many parses. Consider the similar case (5) where a scalar item is embedded under a universal quantifier. There the two parses, with embedded and unembedded EXH are in an entailment relation: (5a) entails (5b), and therefore the latter does not survive the SMH, even though it has been claimed to be a possible reading by some experimental studies (for relevant discussion, see Gotzner and Benz (2018) and references therein).

So, a more permissive variant of the SMH is needed, and one has been formalized by Fox and Spector (2018), as an Economy Constraint (EC). Simplifying to a degree, they propose that:

(13) **Economy Constraint (EC)**
Parses are banned if they contain an occurrence of EXH that is *weakening*, in the sense that removing it would yield a stronger meaning.[5]

[4] The computation of scalar implicatures arising from disjuncts, as in (9) and (10), involves some intricacies; the result we give is what the procedure of Fox (2007) predicts.

[5] They in fact make a distinction between *incrementally weakening* and *globally weakening*, but this is not

Going back to (5), with a scalar item embedded under a universal quantifier, the two parses (5a) and (5b) survive the EC: they are both stronger than other versions of them with fewer EXH operators. Many other cases will remain unchanged, because the reasonings above were about finding stronger meanings by removing embedded EXH operators (that is the case for embedding under negation (8) and disjunctions (9)). However, this EC constraint doesn't rule out (10b) nor (11b) anymore: the original version of the SMH was able to rule these out by comparison with parses with EXH operators in other creative positions, but simply removing the embedded EXH actually results in a weaker meaning, compliant then with the EC.

Finally, note than both the SMH and the EC predict a complete ban on exhaustification embedded below negation, as EXH could only play a weakening role in this position. It isn't entirely clear whether this is correct in at least one case, specifically that of *free choice* inferences. The phrase "free choice" (FC) refers to patterns of inferences like that in (14a), as schematically presented in (14b). Taken literally, the initial utterance should not entail that the addressee is allowed to choose; the fact that it is usually understood that way has been thought to derive from some exhaustification mechanism (e.g., Kratzer and Shimoyama 2002 and Fox 2007).

(14) a. Sue may take an apple or a banana.
$\rightsquigarrow$ Sue may take an apple, Sue may take a banana.
b. $\Diamond(A \vee B) \rightsquigarrow \Diamond A \wedge \Diamond B$

Importantly, the same analysis may also predict an additional, optional inference, that Sue may not take an apple and a banana *at the same time*. This is indeed a possible inference from (14a). We will follow Fox (2007) in calling it the anti-conjunctive (AC) inference and distinguishing it from the FC inference proper, outlined in (14b).

If the FC and AC inferences may be derived below negation, then (15) should have a variety of readings besides its literal reading (16a). First, as paraphrased in (16b), it should have a reading in which FC is derived below negation; a reading true if, for instance, you may take an apple, but you may not take a banana. We think that this reading, while perhaps less natural than the literal one, is available, in the sense that compared to other cases discussed here such as that of (8) with a regular scalar item under negation, we can access it without special intonation or other marking.[6] On the other hand, the reading in (16c), which one would obtain by computing the anti-conjunctive inference below negation, is not accessible at all. Finally, the reading paraphrased in (16d) takes into account both inferences at once, it seems inaccessible as well.

(15) It is not the case that you may take an apple or a banana.

(16) a. Literal reading: you may neither take an apple nor take a banana.
b. Neg{FC}: taking an apple and taking a banana are not both allowed.
c. Neg{AC} (unattested): you may neither take an apple nor take a banana, or you may take both at once.
d. Neg{FC+AC} (unattested): taking an apple is not allowed, or taking a banana is not allowed, or taking both at once is allowed.

relevant to our discussion.

[6]Fox (2007, fn. 16) classifies the weak reading as an instance of an "intrusive" local implicature, and considers that it is marked, but also remarks that this judgement is less clear than in similar cases of embedded exhaustification.

Crucially, all three non-literal readings are constructed by adding inferences below negation. They are therefore all weaker than the literal reading, and should all be equally ruled out by the SMH or the EC, if one accepts an exhaustification analysis for these inferences. If real, the (even relative) availability of the Neg{FC} reading (16b), is thus a challenge for the SMH and the EC which do not seem to deliver the right constraints on the distribution of EXH.

To sum up, the predictions of each constraint are given in Table 3.1. The SMH and EC have different strengths and weaknesses, but they are both based on the idea that parses with and without EXH compete with one another. Instead, we are going to propose a different constraint, the Connected Meaning Hypothesis (CMH), which excludes possible parses on the basis of their output value only. The CMH is thus of a different nature, and it will have a different motivation, as such it could even co-exist with the SMH or EC. Beyond its conceptual motivation, it is a good contender also because, as we will now present, its coverage of the empirical landscape is just as good as for the other options currently available.

Example, description and schematic parse			Fact	SMH	EC	CMH
(5a)	Universal quantifier	$\forall$EXH	YES	YES	YES	YES
(5b)	Universal quantifier	EXH$\forall$	YES	NO	YES	YES
(6a)	Non-monotonic quantifier	QEXH	YES	YES	YES	NO
(6b)	Non-monotonic quantifier	EXHQ	YES	YES	YES	YES
(8)	Negation (basic case)	$\neg$EXH	NO	NO	NO	NO
(9)	3-way disjunction	EXH($A \vee$ EXH($B \vee C$))	NO	NO	NO	NO
(10)	Scalar item in disjunct	$A \vee$ EXH	NO	NO	YES	NO
(11)	Existential quantifier	$\exists$EXH	NO	NO	YES	NO
(15)	Negation Free Choice	$\neg$EXHFC	YES	NO	NO	YES

Table 3.1 – Availability of inserting EXH in various environments, as observed and as predicted by various constraints. SMH: Strongest Meaning Hypothesis (Sauerland 2012); EC: Economy Condition (Fox and Spector 2018); CMH: Connected Meaning Hypothesis (current proposal).

3.2 Proposal: the Connected Meaning Hypothesis

In this section we propose a different principle that may contribute to restrict the distribution of EXH, by ruling out parses unless they produce 'good' outputs. A good output will be one that is 'connected', where connectedness is a formal property that has been proposed first to account for constraints on the meaning of content words, which was then extended to generalized quantifiers, and that we here propose to generalize to full sentential meanings. We will present this constraint and its origins in this section informally, showing how it intuitively covers the cases discussed so far. In the next section, we set up a formal fragment for which we derive theorems closely mimicking the empirical generalizations as they are known so far.

3.2.1 Constraints on words

A body of literature in semantics and related fields has been devoted to finding invariants or constraints that lexical items uphold or mostly uphold. The observation is that many lexical items share some abstract property, and that many other conceivable concepts which miss that property are lexicalized only in few or no languages, thus leading to the conjecture that the space of possible words is constrained by this property (see von Fintel and Matthewson 2008 for a survey of such conjectures, among other things). Horn (1973) makes this observation in the specific case of English logical words, such as connectives and quantifiers. Barwise and Cooper (1981) identify several mathematical properties often verified by lexicalized quantifiers (envisioned as logical functions) across languages. Among those is *monotonicity*. Monotonicity can work in two directions: if a set is part of the denotation of an *increasing monotonic* quantifier, then any larger set also is; if a set is part of the denotation of a *decreasing monotonic* quantifier, then any smaller set also is. For instance, the denotation of "having a pet" is a superset of that of "having a cat". Because "everybody" is an increasing monotonic quantifier, (17a) entails (17b). Because "nobody" is a decreasing monotonic quantifier, (18b) entails (18a). Neither (19a) nor (19b) entails the other, showing that "exactly two people" is not monotonic.

(17) a. Everybody has a cat.
 b. Everybody has a pet.

(18) a. Nobody has a cat.
 b. Nobody has a pet.

(19) a. Exactly two people have a cat.
 b. Exactly two people have a pet.

Even though non-monotonic quantifiers are easily conceivable, Barwise and Cooper (1981) identify a bias in favour of monotonic quantifiers in the lexicon.[7] Katzir and Singh (2013a), without using the word, propose that the meaning of logical words is represented in terms of monotonic primitives, and that non-monotonic items require a complicated representation and are therefore marked.

Chemla, Buccola, and Dautriche (2019) investigated this issue and proposed to study a sister notion of monotonicity, connectedness. They first show that for a quantifier q, monotonicity of q is equivalent to connectedness of q and of its negation $\neg q$. Then, they present observation of lexicons and results from learning experiments arguing for a bias towards connectedness for quantifiers in humans, and even in baboons in Chemla et al. (2018). One further interest of this notion of connectedness is that it actually originated not as a constraint on quantifiers, but in fact as a constraint on the denotation of content words, like nouns, as championed by Gärdenfors (2004) (see also Gärdenfors 2014 for an extension to other types of words). Formally, the denotation of a word is connected if for any two elements in it, any elements that are *in-between* those two are also in the denotation. For instance, if one knows that the word "animal" denotes a connected set of species, and that the duck and the badger are instances of animals, one can deduce that the platypus is an animal. This is, of course, if one thinks that the platypus is in-between ducks and badgers. Hence, in general, the notion of connectedness depends crucially on that of in-betweenness. In the case of content words, one needs an in-betweenness relations over objects, which is

[7] To be precise, they identify a bias in favour of quantifying determiners that produce monotonic quantifiers, such as "every".

often intuitive but certainly difficult to formalize in general; in the case of quantifiers, one needs an in-betweenness relation over sets of objects, and the formal, canonical order of set inclusion has been argued to be productive. We will engage with similar issues once moving to the current enterprise about plain sentential meanings.

3.2.2 Principle: connectedness beyond words

Traditionally, connectedness constraints are meant to explain category learning, and have consequences for word learning: humans naturally group together objects which lie in connected areas of their mental space, and as a result languages end up with connected meanings for content words. The key factor then is a bias towards recognizing connected patterns, which can explain generalizations occurring even with little evidence (e.g. by recovering a smooth and connected category from a few sparse examples). Our proposal builds upon work on concepts and words: we submit that the relevant notion of connectedness, already expanded in the literature from just content words to all words, is an even broader one, and that a similar preference for connected propositions could arise, and should lead to some detectable effects. These effects may not be seen in acquisition, because propositions are not acquired *per se* (although it may indirectly affect the possible meanings of words by making sure that the available inventory will not create connectedness too often). Where they could very well show up, however, is when it comes to disambiguation. One case of a disambiguation problem is that of parsing a sentence potentially containing a silent EXH, and we intend to show that a connectedness constraint makes interesting and empirically testable predictions for this case. Interestingly, in work parallel to ours, Solt and Waldon (2019) put forward a conceptually very similar proposal. The main difference is that their empirical focus is on the interpretation of bare numerals while, as we will see, we study the consequences of connectedness constraints for other cases, such as Free Choice effects.[8]

Since we are concerned with the distribution of EXH, the objects of interest to us are not single words, whether content words or logical words, but entire propositions. We claim that there is a straightforward way to apply the notion of connectedness to propositions: for any two worlds where a connected proposition is true, it is also true in every world *in-between* those two, for some intuitive notion of in-betweenness among worlds. Our approach towards defining in-betweenness on worlds will be detailed in Section 3.4.2; we essentially derive it from an ordering on worlds, where the ordering depends on which answers to the Question under Discussion (QUD) are true at each world.

Our claim, then, is that non-connected meanings are problematic. Consider (20), for instance; it is not an easy sentence to cope with, which we claim could be because it is not connected, in the following sense: take three worlds w_{10}, w_{30} and w_{50} which are most similar to one another except that there are 10, 30 and 50 people attending my class. The world w_{30} intuitively counts as in-between the worlds w_{10} and w_{50}. Yet, the sentence (20) is true at w_{10} and w_{50}, and false in between, at w_{30}. So, (20) is not connected, and indeed it is quite a weird sentence, suggesting that there is a dispreference towards propositions whose meaning isn't connected.

(20) ? Fewer than 20 or more than 40 people attended my class.

[8] Solt and Waldon (2019) correlate the availability of different interpretations of bare numerals (as assessed from a corpus study and experimental results) to a preference for connected meanings, where connectedness is taken to be a property of the set of true answers to certain questions. This is essentially a special case of the definition we are going to adopt in Section 3.4.2.

Of course, (20) is not an impossible sentence. It is certainly not a plainly deviant sentence *per se* and non-connectedness may not be problematic in general (more on this below). But we propose that EXH, which is optional, should not be recruited if it ends up creating non-connectedness:

(21) **Connected Meaning Hypothesis (CMH)**
Among the parses of a sentence (with or without EXH), those that result in non-connected meanings are dispreferred/marked.

This first statement of the principle remains at a non-formal, intuitive level. We will resolve this in backward order. Only Section 3.4 will make all of the details explicit: it will include the derivation of an in-betweenness relation and its associated connectedness property that account for the examples to be reviewed here, and a formal fragment with theorems as to what types of parses are penalized by this principle (e.g., exhaustification under negation will be generally penalized). In Section 3.3, the status of this constraint will be discussed, for at this point it may be implemented as a strict impossibility constraint, a contextual constraint, or an entirely soft constraint. But for now, in the upcoming Section 3.2.3, we want to show how the notion applies intuitively and how it captures the essential facts about the distribution of EXH presented earlier. In short, we are going to argue that the unattested parses typically have weird paraphrases, weird in some intuitive sense related to non-connectedness.

3.2.3 Application

Consider first the paradigmatic example of blocking EXH from under negation, (8), repeated below: the parse in (22), and the meaning it ought to result in, are intuitively unavailable. The CMH proposes to account for this by showing that this meaning is problematic. To put it simply, there is something complicated about the idea that John did either none or all of the homework. Concretely, this meaning is non-connected, in the same way as the meaning of (20) is non-connected. To see this, consider the sequence of worlds w_a, w_b and w_c, which are as similar as possible except that: in world w_a, John spent the day watching series, in world w_b, he took the time to do half of his homework, in world w_c, he worked hard to complete his homework. Intuitively, w_b is in-between w_a and w_c, but (22) is true at both w_a and w_c, and false at w_b. This is schematized in (23), where we see an F in between two Ts.

(22) Sentence: It's not true that John did some of his homework.
Unattested parse: *It's not true that John did EXH [some of his homework].
Intended meaning: It isn't true that John did some but not all of his homework.

(23) Offending sequence of worlds:

	¬EXH(A)	EXH(¬A)
World w_a: John did no homework	T	T
World w_b: John did half of his homework	F	F
World w_c: John did all of his homework	T	F

As schematized in (23) as well, we see that this sequence is not a problem for the parse without EXH, which in fact is fully connected: the sentence is true whenever John did none of his homework, and a world in-between two worlds where he didn't work ought to be one where he didn't work either. Hence the CMH predicts the interpretation without EXH to always be favoured over the one with EXH under negation.

Consider now (9), repeated here as (24). It has the structure $A \vee B \vee C$. (24) can be parsed in several ways, including as (24a), and as (24b). As before, we claim that the meaning of (24b) is hard to reason about, even based on the explicit paraphrase. It is also non-connected in the following intuitive sense: take a world where John meets only Ann, a world where he meets Ann and Carol, and a world where he meets all three people. (24b) is true at the first and third world, but not at the second one, which is intuitively in-between the other two. This is summarized in (25).

(24) John will meet Ann, Bill or Carol.

a. Attested parse: $\text{EXH}[A \vee B \vee C]$
Meaning: John will meet exactly one person within Ann, Bill and Carol.

b. Unattested parse: $^*\text{EXH}[A \vee \text{EXH}[B \vee C]]$.
Intended meaning: John will meet exactly one person within Ann, Bill and Carol, or all of them.

(25) Offending sequence of worlds:

	$\text{EXH}(A \vee \text{EXH}(B \vee C))$	$\text{EXH}(A \vee B \vee C)$
World a: A	T	T
World b: A and B	F	F
World c: A and B and C	T	F

As before, we see that this sequence of worlds is not a counter-example to the connectedness of the attested parse (24a), and in fact, we argue that (24a) is connected. First, for any two worlds where John meets just one person, the same in both, he certainly also meets them in any world in-between. For a world where he meets just Ann and a world where he meets just Carol, one could imagine that in some in-between worlds he meets both or neither, but we adopt a notion of in-betweenness where there is no world in-between those two, because there is no obvious way to relate one to the other (cf. Appendix 3.A). This makes the meaning of (24a) connected, explaining why it is felicitous while (24b) isn't.

Let's turn to (11), repeated here as (26a). The structure of (26a) can be represented by the formula in (26b).

(26) a. There are employees who saw some of the criminals.

b. $\exists e, \exists c, S(e,c)$

We will assume that both elements with existential force ('there are' and 'some') have a universal quantifier ("all of the...") as an alternative. Several parses can be considered. In (27) we represent some of them, starting with those in which EXH always take into account all alternatives potentially triggered from its scope, as signalled by the indices that match with the set of scalar items in the scope of EXH in each case. For completeness we also present there two other options that can be treated similarly.

(27) a. $^*\ \exists_1 e, \text{EXH}_2 \exists_2 c, S(e,c)$
Some employees saw some but not all criminals.

b. $\text{EXH}_{1,2} \exists_1 e, \exists_2 c, S(e,c)$
Some but not all employees saw some criminals, and no employee saw all criminals.

c. $^*\ \text{EXH}_1 \exists_1 e, \text{EXH}_2 \exists_2 c, S(e,c)$
Some but not all employees saw some but not all criminals.

d. $\text{EXH}_2 \exists_1 e, \exists_2 c, S(e,c)$

Some employees saw some criminals, but no employee saw all criminals.

In (28), we use the same method as before to show that (27a) (as well as (27c)) is not connected. This counter-example is compatible with the connectedness of (27b) (and (27d)), however. Proving connectedness is more demanding: one needs to show that *no* counter-example could be found, which we submit is the case here for reasonably intuitive notions of in-betweenness (as is made explicit in Appendix 3.A). Overall then, the CMH appropriately predicts that (27a) and (27c) should be unavailable, and that (27d) and (27b) are possible parses.

(28) Offending sequence of worlds:
Consider that there were a couple of criminals (so, seeing one criminal counts as seeing some but not all), and that among the employees none has seen any criminal, except for A and B, for which the situation is as follows:

	(27a)	(27b)	(27c)	(27d)
World a: *A* saw one criminal, *B* saw none	T	T	T	T
World b: *A* saw all criminals, *B* saw none	F	F	F	F
World c: *A* saw all criminals, *B* saw one	T	F	T	F

Consider now (5), repeated as (29a). Again, we give it a formula representation in (29b). There are two ways to parse (29a) with EXH, listed in (30).[9]

(29) a. Every student did some of the homework.
b. $\forall s, \exists h, D(s,h)$

(30) a. $\forall s$, EXH$\exists h, D(s,h)$
Each student did a proper part of their homework.
b. EXH$\forall s, \exists h, D(s,h)$
Every student did some of the homework, but some students didn't do all of it.

We think both of those have a connected meaning. The fact that (30a) is connected is quite intuitive: in-between two worlds where (30a) is true, the extent of their homework each student did ought to be in-between the extent of their homework they did in the extreme worlds. If it is a proper part on both sides, it is a proper part in-between. We can also give a sense as to why (30b) is connected: consider two such worlds w_a and w_c where (30b) is true, that is, in both of these worlds, all students did part of their homework. For a world w_b to be in between w_a and w_c, under our definition, all students would have to have done more homework in w_b than in one of these worlds (say, w_a), and less than in the other world (w_c) (cf. Appendix 3.A). Necessarily then, because not all students have done part of their homework in w_c, this would also hold in w_b, making (30b) true at w_b. Overall, then, both (30a) and (30b) are connected, and the CMH therefore does not exclude these parses.

As the next example, we will look at (6), repeated here as (31), and consider the two parses in (32).

(31) Exactly three employees saw some of the criminals.

(32) a. EXH [Exactly three employees saw some of the criminals.]
Meaning: Exactly three employees saw some of the criminals, less than that saw all of them.

[9] We ignore the possibility of obtaining alternatives by replacing the universal quantifier at the top of the sentence with an existential quantifier.

b. Exactly three employees saw EXH [some of the criminals.]
Meaning: Exactly three employees saw some but not all of the criminals.

As seen in the sequence of worlds in (33), (32b) doesn't have a connected meaning. Meanwhile, it can be shown that (32a) does under identifiable assumptions (cf. Appendix 3.A). Thus, the CMH predicts (32b) to be unavailable. While (32a) does seem to be the more accessible reading, it isn't clear that (32b) is as inaccessible, and there is experimental evidence for its existence. We will discuss limitations of the CMH and potential reasons for its violation below.

(33) Offending sequence of worlds:
Again, assume that there were a couple of criminals (so, seeing one criminal counts as seeing some but not all). Among the employees none has seen any criminal, except for A, B, C, D, E and F for which the situation is as follows:

	(32a)	(32b)
World a: A, B, C saw one criminal, D, E, F saw none	T	T
World b: A, B, C saw all criminals, D, E, F saw none	T	F
World c: A, B, C saw all criminals, D, E, F saw one criminal	F	T

Our final examples involve so-called Free Choice effects. In particular, consider the case of Free Choice under negation, as in (15), repeated here as (34), with its schematic structure also given there. As we already discussed, some analyses predict that $\Diamond(A \vee B)$, when subject to exhaustification, may become $\Diamond A \wedge \Diamond B$. Thus, if (34) is parsed with an embedded EXH as in (35a), it ought to be equivalent to (35b) (this is the Neg{FC} reading from (16b)).

(34) It is not the case that you may take an apple or a banana.

(35) a. Possible parse: $\neg$EXH$\Diamond(A \vee B)$
b. Meaning: $(\neg\Diamond A) \vee (\neg\Diamond B)$
(You may not take an apple or you may not take a banana.)

Despite the fact that it requires an EXH operator under negation, the meaning in (35b) is connected, and therefore complies with the CMH. To get an intuition as to why this is true, take two worlds where (35b) is true. There are three possibilities: if in both worlds you may not take a banana, then it is true in any in-between world; if in both worlds you may not take an apple, then again it should hold in any in-between world; finally, if in one world you may take an apple but not a banana, and in the other it is the opposite, then there is no obvious way to relate the two worlds, and by our definition there is no world in-between. Thus, the CMH tolerates Free Choice inferences under negation, what we have called Neg{FC} above in (16).

One can go further and investigate the other options mentioned in (16), involving the anti-conjunctive inference: neither the Neg{AC} ((16c)) nor the Neg{FC+AC} ((16d)) readings are connected, and they are therefore appropriately banned by the CMH. We leave the details to the reader and to Table (36), which provides a counter-example valid for both cases. Note that the CMH makes no difference between the Neg{FC} reading and the literal reading, as they are both connected; thus the CMH does not alone provide an explanation as to why the literal reading is usually the first one speakers access. It may be that other principles are needed *in addition* to the CMH to explain the full range of preferences within the readings left possible by the CMH.[10]

[10] Interestingly, one could adopt some form of the SMH for that purpose — recall that the SMH was initially

(36) Offending sequence of worlds:

		(16c)/(16d)
World a:	There is nothing you may take.	T
World b:	You may take an apple, you may take a banana, but not both at once.	F
World c:	You may take both (at once).	T

Finally, as mentioned by a reviewer, a similar pattern to FC is that of *distributive inferences*, which are treated as a kind of FC by certain theories. Consider in particular the dual case in which a disjunction is embedded under a universal modal (and not an existential modal), and then under negation:

(37) Mary is not required to take Syntax or Logic.
⇝ Mary is allowed to take Syntax and she is allowed to take Logic.

This gives rise to a so-called distributive inference of the form $\neg\Box(S \vee L) \rightsquigarrow (\neg\Box(S \vee L)) \wedge \diamond S \wedge \diamond L$. The enriched meaning is a conjunction of terms (including the literal meaning), where every term is monotonic. As we are going to show, this entails that it's connected per Theorem (55). Thus the CMH doesn't rule out this inference. Note that as a conjunction of terms where one is the literal meaning, the enriched meaning is also stronger than the literal meaning: thus the SMH and EC do not rule it out either, correctly so. The CMH also agrees with the SMH and EC in banning a parse of (37) where EXH occurs below negation. Such a parse would have a meaning that can be paraphrased as "either Mary is not required to take Syntax nor Logic, or she is required to take Syntax, or she is required to take Logic"; that meaning is non-connected for roughly the same reasons as (16d). We think that this reading is indeed unavailable. To conclude, the CMH makes correct predictions for both universal and existential modals, when they host a disjunction and are embedded under negation.

3.3 Discussion

Having established the potential of the CMH to be at the source of a range of empirical facts, we would now like to discuss the conceptual motivations of the CMH, and explore some of the parameters we have currently left unset: is the CMH a strict or a soft constraint? what is its status as a linguistic constraint and how does it interact with other linguistic constraints?

3.3.1 Status and origin of the CMH

To get started, one may ask whether the CMH is a hard constraint, as in: non-connected meanings are ungrammatical. This is not the spirit of the CMH, nor does it seem to comply with the introspective status of the relevant sentences. Looking back at the case with no EXH operator, it would be strange to declare a sentence like (20) ungrammatical; intuitively, it is merely an unlikely thing to say, something hard to think about even, which is what the CMH tries to capture. The idea then is that the CMH prevents us from using roundabout ways, the application of EXH, to derive meanings we will end up not being comfortable with.

introduced to account for other phenomena than the distribution of EXH, and therefore, even if we explain the distribution of EXH through the CMH, we might expect the SMH to still be active in disambiguating between available readings, whether or not those readings have been filtered by a principle such as the CMH.

One may expect from this general approach that the CMH should not be a strict constraint: even if we may have difficulties manipulating non-connected objects, we ought to have ways to do it, in case this happens to be useful. So, even if non-connectedness will be restricted, general reasoning abilities and language should find a way to allow us to manipulate non-connectedness. In fact, language surely does provide the means to express non-connected meanings, often times through disjunction: after all, (20) is hard, but we can construct it and work with it. Other sources of non-connectedness in natural languages involve embedding under non-monotonic constituents, as in (38) or (39); again, these examples are hard and may require a second thought, but they are possible sentences.

(38) It isn't the case that John did exactly two exercises.[11]

(39) Some students did exactly two exercises.

So, the CMH is best viewed as a the result of a cognitive dislike for non-connected meanings, which results in the (soft) blocking of some parses that would unnecessarily create such non-connected meanings. Ideally, one should try to explain *why* we are not comfortable with non-connected meanings, and show that this cognitive distaste for non-connectedness generalizes beyond the linguistic domain (as attempted to in Chemla, Buccola, and Dautriche 2019 and Chemla et al. 2018). For now, it is sufficient to take away that the CMH is not to be thought about as a hard constraint. We will thus now discuss several potential sources for CMH violations.

3.3.2 Contextually salient non-connectedness

If we take the CMH to derive from a cognitive bias against non-connectedness, we might expect that context could help alleviate the difficulty, making some non-connectedness proposition cognitively easy or salient, to the point that violations of the CMH could be saved. This view is somewhat risky, in that it could deprive the CMH of much of its predictive power. However, if one had an independent measure of the difficulty of a particular non-connected meaning in a particular context, the predictive power could be recovered, with the expectation that CMH violations would be higher, for difficult non-connected meanings. At the very least, we can start investigating the issue by trying very clear cases of easy non-connectedness.

Consider for instance (40a). By default, it seems to have a connected meaning, whereby some students did two or more exercises. But in a context as the one given in (40b), which makes salient and important the possibility that a student does exactly two exercises, it may receive a non-connected interpretation, with the numeral receiving an 'exact' reading under the existential quantifier, therefore obtaining the same non-connectedness as (39) (or (27a)).

(40) a. Some students did two exercises.
b. Supportive context: if a student does exactly two exercises, the grading database crashes.

Thus, the availability of the non-connected meaning in (40) could be taken to indicate that the CMH's strictness is only relative to the difficulty of conceiving of non-connectedness

[11]This is one of several cases in this article where finding examples involving matrix negation is made difficult by structural ambiguities. Here, matrix negation makes very salient the reading that there are exactly two exercises John didn't do.

in the context.[12] But it is most important to act with caution here because it does not seem to be so easy to allow violations of the CMH by simple manipulations of the contexts in other cases. For instance, embedded exhaustification isn't as clearly accessible in (41) as in the very similar (40a), that is, with "some" as a scalar item, rather than a numeral.

(41) Context: students who did nothing get a C, those who did part of the work get a B, those who did all of it get an A. If a student gets a B the database crashes.
Some students did some of the exercises (? and therefore the professor couldn't enter their grades).

Hence, the previous result about (40) may be specific to numerals since, as already seen from (38) or (39), non-connected propositions are particularly easy to obtain from numerical expressions. There would be several ways to explain such a fact. First, (40) is only a violation of the CMH if one assumes that the non-monotonic reading of "two" derives from exhaustification. If that isn't the case, we do not have such a clear case for the sensitivity of the CMH to contextual salience in the first place.

Second, it could be easier to reason about non-monotonic or non-connected things in numeric or mathematical contexts. Intuitively, the concept of "exactly two" is more accessible than that of "a proper part". As a result, EXH remains very attractive despite the threat of the CMH. In this latter case, then, the CMH does show sensitivity to content, rather than to context. The resulting view would be highly compatible with the idea that the distribution of EXH is governed by the shape of its output, complemented with the idea that the mathematical domain has a pragmatics of its own.

3.3.3 Contextually connected non-connectedness

Another way in which context could interact with the CMH is through the common ground, considering that it could transform non-connected meanings into contextually connected meanings. So, a particular proposition may be non-connected, while the intersection of it with the common ground may be connected. Assume then a notion of logical connectedness as well as a notion of contextual connectedness. If the CMH is a principle that bans meanings that are hard to think about, and if the common ground can be incorporated without efforts into the decoding of linguistic meanings, then we may find that contextual connectedness is the appropriate notion, to the effect that the common ground is sufficient to relieve the burden of non-connectedness.

Unfortunately, the current evidence is hard to judge. Consider for instance (42). In this context, the sentence is either trivially true (a contextual tautology) with its normal con-

[12] A similar, but more abstract effect can perhaps be found in the experimental results of Chemla and Spector (2011). As we already mentioned, they report that participants in a judgement task seemed to access parses similar to (6a), repeated here as (i), where EXH is embedded below a non-monotonic quantifier. This is problematic for the CMH, as this parse doesn't result in a connected meaning. Interestingly, they also report that in a second judgement task, the same participants seemed to access parses where EXH occurs below negation more easily than other participants whose first task did not involve non-connected meanings.

(i) Exactly three employees saw EXH [some of the criminals.]

Chemla and Spector (2011) speculate that the first category of participants were somehow primed on embedded exhaustification, because something about their first task made such parses more salient. We may venture that, while we cannot explain why these participants were able to access non-connected readings in the first place, their first task accustomed them to non-connectedness and let them relax the CMH and access non-connected meanings in the second task too.

nected meaning, or it is contextually connected (but not logically connected). This is thus an ideal situation to see the otherwise unattested meaning rise. It is unclear however what the facts really are: the sentence may well receive that interpretation, but it still sounds a bit deviant.

(42) Context: You have to do all exercises to get an A. All students did part of the homework.
Some students did some of the exercises (? and therefore didn't get an A).

Thus, there is no evidence that contextual connectedness, rather than bare connectedness, has much to offer.[13]

---

[13]Interestingly, an effect of the CMH may be seen in the interaction between the common ground and *overt* strategies for exhaustification. Both unusual focus marking and "only" adverbs have effects similar to covert exhaustification: the following two sentences both unambiguously imply that John didn't do all the relevant exercises. This may be analysed by taking "only" to have similar semantics to EXH, and focus marking to co-occur with EXH.

(i) a. John only did some of the exercises.
b. John did SOME of the exercises.

As a result, just like unmarked occurrences of EXH, these constructs introduce the possibility of non-connectedness when embedded in some environments. However, a closer look suggests that even in such cases, non-connectedness is avoided. Consider what happens when these overt exhaustification strategies are embedded in the problematic negative or existential environments:

(ii) a. It's not the case that John only did some of the exercises.
b. It's not the case that John did SOME of the exercises.

(iii) a. Some students only did some of the exercises.
b. Some students did SOME of the exercises.

The overall meanings obtained actually qualify as connected: in (ii), the sentences' overall meaning is that John did all of the exercises (and not that he did either none or all of the exercises). This is diagnosed by the fact that (iva) is a felicitous continuation of both examples in (ii), while (ivb) is not. Similarly, in (iii), the sentences' overall meaning seems to be that all students did some of the exercises, but not all students did all of the exercises.

(iv) a. He did all of them.
b. He did none of them.

A common analysis of these 'overall meanings' states that they involve not only assertive components, but also presuppositions or some form of non at-issue content (see Coppock and Beaver 2014 or Roberts 2006 for discussion about "only", Schwarzschild 1999 about focus marking), projecting out of the environments (through universal projection in the case of the existential environments *à la* Heim 1983, or more recently Schlenker 2008, *pace* Beaver 2001). These special parts of meanings turn out to block non-connectedness. To put it simply, one may say that the presuppositions of "only" and of unusual focus marking protect them from introducing non-connectedness when embedded in these environments.

We could thus argue that connectedness constraints have a role to play not only in the selection of the meanings of lexical items, but also in the selection of their presuppositions. One may then ask why EXH does not grow a presupposition that would have the same protective effect: a possible answer is that interpreting a sentence with an unmarked EXH is always optional, and a parse without EXH can always be selected instead if non-connectedness would occur otherwise. Another question is why certain other items that, like "only", introduce non-monotonicity (such as non-monotonic quantifiers) do not acquire a similar protective presupposition. A language with a certain level of expressive power would certainly need to include items capable of resisting all sorts of non-connectedness constraints, but it is an open question to tell which items may be elected to do so.

3.3.4 Interaction of the CMH with Hurford's constraint

The CMH is meant to block some occurrences of EXH. One may ask what happens when the satisfaction of other constraints may force the occurrence of EXH. Chierchia, Fox, and Spector (2011) present a variety of examples where Hurford's constraint has precisely this effect; those examples are (specific cases of) *Hurford disjunctions*.

Recall that in a Hurford disjunction, the exhaustification of a disjunct is necessary to avoid that disjunct to be redundant. This derives from a principle called Hurford's constraint (HC): HC states that disjunction where the first disjunct entails the second are infelicitous (see e.g., Katzir and Singh 2013b, Mayr and Romoli 2016). We can construct an example where HC forces local exhaustification, but where this results in a non-connected meaning. This is the case for instance of (43). Hurford's constraint pushes in favour of the insertion of EXH in the first disjunct, but the resulting meaning would be non-connected as long as there are more than 3 problems, and we allow 'three problems' to be an available alternative available for the exhaustification mechanism. The non-connectedness of the resulting meaning can be established from (44). We observe that the sentence is acceptable despite the violation of the CMH.

(43) Peter either solved both the first and the second problem or all of the problems. (Chierchia, Fox, and Spector 2011)

(44) Offending sequence of worlds:
Assume there are 5 problems.

	(43)
World a: Peter solved problems 1 and 2.	T
World b: Peter solved problems 1 through 3.	F
World c: Peter solved problems 1 through 5.	T

Overall, then, it seems that the CMH is liable to being overridden by other principles that force the presence of EXH in certain positions, such as Hurford's constraint. The situation can be described by saying that one should not create connectedness with embedded EXH for no reason, but Hurford's Constraint is actually a good reason. Thus, the CMH appears to be a relatively soft constraint.

3.3.5 The CMH on intermediate constituents

So far, we have only discussed examples where connectedness was evaluated at the sentence level. However, we can imagine that the constraint could be active at any intermediate steps during the construction of the meaning of each constituent. Practically, the connectedness constraint may block occurrences of EXH that result in non-connected meanings not only at the sentential level, but also if it requires going through a non-connected meaning for an intermediate constituent of the sentence (we thank a reviewer for urging us to include this discussion and providing relevant examples).

Our statement of the CMH does not specify whether connectedness applies to proper constituents of the sentence. In this section, we evaluate this possibility by considering two types of empirical evidence. The results go in opposite directions and both examples seem amenable to alternative explanations, however, so we will leave this parameter of the proposal open.

First, consider cases in which the overall sentence is connected, but an exhaustified constituent clause of it isn't. Examples of this kind are in fact discussed by Chierchia, Fox, and Spector (2011). In (45a) (repeated from (4a)) the insertion of EXH somewhere in the first relative clause is necessary for the sentence not to be a contradiction. It can then be verified that the meaning of the first conjunct is non-connected, and that the overall sentence has a connected meaning.

(45) a. Every professor who fails most of the students will be assigned a new class next year, and every professor who fails all of the students will be fired.
b. Parse: Every professor who fails EXH [most of the students] will be assigned a new class next year, and every professor who fails all of the students will be fired.

One could thus take the felicity of this example as evidence that the CMH does not affect proper constituents, and that only the overall meaning matters. However, since the parse without EXH results in a contradiction, we could also say that a constraint favouring non-contradictory parses takes precedence over the CMH here, much like we noted that Hurford's constraint takes precedence over the CMH. Under this line of explanation then, this example remains compatible with applying the CMH to intermediate constituents.

The second category of cases are those in which the overall sentence is non-connected, whether EXH is present or not. We may ask whether local insertions of EXH could be banned because they would create non-connected local constituents. In (46a), the disjunctive quantification results in a non-connected meaning even for a parse with no EXH. The current statement of the CMH bans EXH below negation in simple cases, but it does not predict a contrast between the parse in (46b) and a parse without EXH: both violate the CMH. Intuitively, (46a) is hard to process, but it is nevertheless interpretable and its potential interpretations show a contrast: a reading corresponding with a parse without EXH seems more accessible than one with EXH as in (46b). A similar example is given in (47a).[14]

(46) a. ? Fewer than 20 or more than 40 people didn't do some of their homework.
b. Possible parse: Fewer than 20 or more than 40 people didn't EXH [do some of their homework.]

(47) a. An even number of students did some of the homework.
b. Possible parse: An even number of students did EXH [some of the homework.]

This time, we might take (46a) as evidence that the CMH applies to intermediate constituents and not just to entire sentences. Then, we can readily explain that the contrast persists for (46a): having EXH below negation will create a non-connected constituent, no matter how we embed it further. However, (46a) may also be amenable to other explanations: in particular, one could to define a global, graded notion of connectedness, such that a meaning with, intuitively, more gaps in its extension would be less connected. With such a definition,[15] the parse in (46b) may be *less* connected than alternative parses. A global (and

[14]We prefer to focus on (46a), however, as it may be that mathematical expressions such as "an even number" prompt us to consider specific conceptual spaces with a different topology where our intuitions of connectedness for numbers break down. In this instance, it could be that the relevant order on groups of students for (47a) is the partial order deriving from a polar QUD "even or not?", rather than the one deriving from the usual order on cardinalities. This makes connectedness harder to assess in mathematical contexts.

[15]It is feasible to generalize our notion of connectedness into something graded. For instance, as a proof of concept, we could say that a proposition p is k-unconnected if there is a chain of k p-worlds such that there is a non-p-world in-between every consecutive world in the chain. Then, our non-graded notion of non-

graded) notion of connectedness could therefore explain that (46b) is dispreferred.

3.3.6 The CMH on unembedded exhaustification?

Finally, while we have been focussing on the consequences of the principle on the occurrence of EXH in embedded positions, it is conceivable that even regular, unembedded exhaustification could be affected by the CMH. Indeed, certain choices of non-monotonic alternatives could lead to non-connectedness from global exhaustification alone. In this respect the CMH differs from the SMH and related hypotheses: global occurrences of EXH may only result in strengthening by definition, and will never be banned by the SMH.

As an attempt to find relevant examples, consider (48). In principle, the first statement could be used as an alternative to the second statement, which would then be exhaustified as described at the bottom of the example. This does not seem to happen, however, and the CMH predicts this impossibility since it would result in a non-connected meaning. However, such type of exhaustification may not be that easy and could be blocked for independent reasons: a similar type of exhaustification also does not happen in (49), even though there the final meaning would be connected. Now, there is an independent reason why this reading does not emerge in (49), which is that there was a much simpler way to get at it, not relying on exhaustification, and simply putting in the right number in the sentence in the first place; thus simply replacing 30 with 20 in the target sentence would do the trick. If this is the reason why that reading is not available below, then we need an explanation as to why the reading discussed in (48) is not available, and the CMH could provide one.

(48) Yesterday, there were between 20 and 30 people.
Today, there are between 10 and 40 people.
* $\rightsquigarrow$ There are between 10 and 20 or between 30 and 40 people.

(49) Yesterday, there were between 20 and 30 people.
Today, there are between 10 and 30 people.
* $\rightsquigarrow$ There are between 10 and 20 people.

3.4 The CMH in an explicit fragment

In this section, we present the CMH in a more formal and systematic manner, by investigating definitions and generalizations in a well-defined fragment roughly corresponding to predicate logic, with exhaustification. In Section 3.4.1 we define an explicit fragment capable of mimicking almost all of the cases reviewed so far. In Sections 3.4.2 and 3.4.3 we define monotonicity and connectedness as formal properties of linguistic objects, and prove various theorems about which structures lead to monotonic or connected meanings. Section 3.4.4 connect these results to the empirical generalizations presented above.

3.4.1 Fragment

Our fragment is essentially that of first-order logic, with constants, variables, predicates, negation, disjunction and conjunction, existential and universal quantifiers, all with their

connectedness is equivalent to 2-unconnectedness. The parse of (46a) without EXH is only 2-unconnected, while the parse with EXH is k-unconnected for some large k. This corresponds to the intuition that the parse with EXH has a less connected meaning than the parse without EXH, in the sense that it picks out a more scattered set of situations. We do not explore this possibility any further here.

usual semantics. We need to add the operator EXH to such a fragment. We do so first by adding the straightforward syntactic inductive rule:

(50) If ϕ is a sentence, EXHϕ is a sentence.

The semantics of EXH is as follows, where we assume the existence of an operator ALT which applies to every sentence to provide a set of alternative sentences (we do not give further compositional structure to the formation of alternatives here, see Katzir 2007):

(51) $[\![\text{EXH}\phi]\!] = [\![\phi]\!] \wedge \left(\bigwedge_{\phi' \in \text{ALT}(\phi)} \neg[\![\phi']\!]\right)$

Of importance to us is the fact the semantic value of a sentence is a function from situations to truth value (i.e., it has type $\langle s, t\rangle$). A situation is an interpretation function for the variables and predicates: it maps variables to individuals in some domain and predicates to predicates with the appropriate arity over the same domain. In what follows we will use two notational shortcuts: (i) we will drop double brackets when it is possible to do so without creating ambiguities, (ii) we will write down semantic objects of type $\langle \alpha, t\rangle$ in set notation.

3.4.2 Assumptions and formal definition of connectedness

To continue, the semantic objects must be equipped with more structure. In particular, to define connectedness we need in-between relations for input types and orders for output types — here we will derive the in-betweenness relation for input types from an order, in the same way as Chemla, Buccola, and Dautriche (2019) do.[16]

First, we assume that truth-values are ordered in the usual, conventional sense: $1 > 0$, truth is ranked higher than falsity. From there, we can derive an order for functions of any type ending in t.

Second, we also assume that objects of type s are ordered. Although this would be desirable (see for instance discussion in the literature on counterfactuals, as in Stalnaker 1968; Lewis 1973; Lewis 1981; Kratzer 1979; Kratzer 1981), there is no universal formal way to define an order on situations or possible worlds. However, we claim that a consensual order on worlds can be found in every example of relevance, on a case-by-case basis. Furthermore, we note that orders can be derived from the more familiar, although maybe not more deterministically given, notion of a Question Under Discussion (QUD, see Roberts 1996). A QUD is, in short, a question to which the sentence is understood to be an answer to. It is typically meant to be inferred from contextual information as well as from the utterance itself. Assuming that a QUD is given then, we can map each possible world onto the set of answers to the QUD that are true at this world (the Karttunen denotation of the QUD at this world). A pre-order can then be derived over worlds: any two worlds can be pre-ordered in the same way as the way their corresponding sets of answers are ordered by inclusion. If the situations correspond to the classes of worlds which behave similarly with respect to the QUD, then the pre-order over worlds defines an order over situations. We hope that this connection with the existing literature makes the relevant notion of an order more palpable.[17]

[16] For discussion of potential alternatives to our formal definition of connectedness, see Appendix 3.B.

[17] As pointed out by Benjamin Spector (p.c.), the order on worlds one would derive that way would be essentially the same one that certain theories of exhaustification already require; see for instance Spector (2006), van Rooij and Schulz (2004) and Schulz and van Rooij (2006). Instead of defining it in terms of sets of true answers to the QUD, they use sets of true alternatives to the utterance.

Thus, sentences are assumed to denote functions from an ordered type to another one, and for any such function over ordered types, we can define monotonicity and connectedness (see Appendix 3.B for alternative ways to define connectedness):

(52) **Definition:** Monotonicity and connectedness.
Consider a sentence ϕ of type $\langle \alpha, \beta \rangle$, where α and β are ordered.

a. ϕ is *increasing monotonic* iff for all x, x' of type α, if $x \leq x'$, then $\phi[x] \leq \phi[x']$.
b. ϕ is *decreasing monotonic* iff for all x, x' of type α, if $x \leq x'$, then $\phi[x'] \leq \phi[x]$.
c. ϕ is *monotonic* iff it is increasing monotonic or decreasing monotonic.
d. ϕ is *connected* iff for all x, x', x'' of type α, if they form a *chain*, i.e. $x \leq x' \leq x''$, then:

$$\phi[x'] \leq \phi[x] \wedge \phi[x''] \qquad (*)$$

To tie this together with the informal notion of connectedness we have been using so far, let's define in-betweenness as follows: x' is in between x and x'' iff x, x', x'' form either a chain ($x \leq x' \leq x''$) or an antichain ($x \geq x' \geq x''$). If x' is in-between x and x'' either (x, x', x'') or (x'', x', x) form a chain; in both cases, $(*)$ holds if ϕ is connected. Given that in the specific case where $\beta = t$, $(*)$ can be read as "if $\phi[x]$ and $\phi[x'']$ are both true, so is $\phi[x']$," our old informal definition is just a special case of the new formal one for $\beta = t$.

We can finally mention the following relation between monotonicity and connectedness:

(53) **Result:** Monotonicity entails connectedness.
If ϕ is monotonic, ϕ is connected.

Proof. Let ϕ be a monotonic function, which we will assume without loss of generality to be increasing, and let $x \leq x' \leq x''$ be a chain. Then: $\phi[x''] \geq \phi[x'] \geq \phi[x]$, and: $\phi[x] \wedge \phi[x''] = \phi[x]$. Thus $(*)$ holds, and since this is true for any chain, ϕ is connected. □

For concreteness, Figure 3.1 provides a visual illustration of the possible types of predicates obtained when $\beta = t$, and α is a domain of 4 elements ordered (partially) in a particular way.

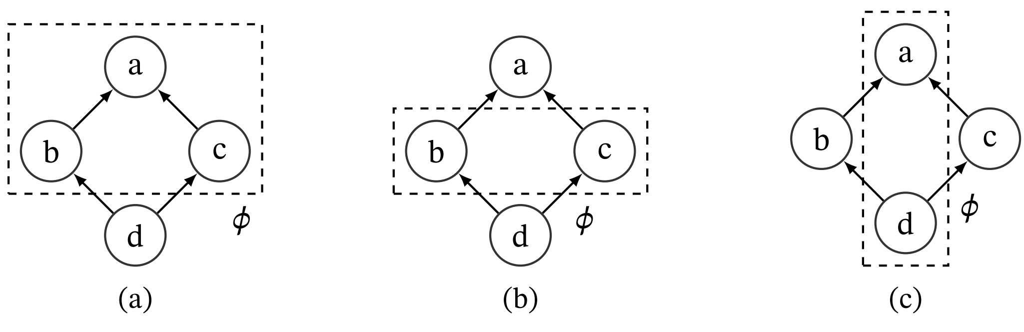

Figure 3.1 – Consider that the domain α contains four elements, ordered as in the diagrams above: a is above b and c, which are incomparable with one another, but both being above d. The dotted boxes provide examples of (a) an upward monotonic and connected predicate (b) a non-monotonic but connected predicate, and (c) a non-connected predicate.

3.4.3 Structures leading to connected meanings

In this section, we ask for several types of structures whether we can, or cannot, be sure that the result will be connected. We inspect the cases following the syntactic construction of our fragment, asking the same question for sentences headed by a negation, a conjunction, a disjunction, a quantifier or, crucially, an EXH operator. In each case, we ask whether the obtained sentence is connected or monotonic, depending on whether the elementary pieces it is made of are monotonic (in the (a) parts of each result) or connected (in the (b) parts of each result). Depending on the cases, we may obtain two types of generalizations: theorems showing that certain structures will necessarily output a monotonic or a connected meaning, and theorems showing that some constructions will not necessarily output a monotonic or a connected meaning.

(54) **Result:** Negative sentences

a. If ϕ is monotonic, $\neg\phi$ is connected.
b. If ϕ is connected, $\neg\phi$ isn't necessarily connected.

Proof. The result in (a) follows as a corollary of result (53): if ϕ is monotonic, $\neg\phi$ is also monotonic, and therefore it is connected.

The following counter-example proves (b): Suppose that three elements form a chain $x_1 < x_2 < x_3$. Define ϕ as the function over these three elements, which is true for x_2 and no other element. Then ϕ is connected, but $\neg\phi$ isn't connected, since its denotation includes x_1 and x_3 but not the in-between point x_2. □

(55) **Result:** Conjunctive sentences

a. If f_i is a collection of monotonic functions, all increasing or all decreasing, $\bigwedge_i f_i$ is monotonic in the same direction.
b. If f_i is a collection of connected functions, $\bigwedge_i f_i$ is connected.

(56) **Corollary:** Universally quantified sentences

a. If ϕ has a free variable y, and if for all individuals e_0, $\phi[y \leftarrow e_0]$ is monotonic in the same direction, then $\forall y, \phi$ is monotonic in this direction.[18]
b. If ϕ has a free variable y, and if for all individuals e_0, $\phi[y \leftarrow e_0]$ is connected, then $\forall y, \phi$ is connected.

Proof. [Proof of (55)] (a) is an immediate consequence of the fact that conjunction preserves the order, i.e. that it is itself an upward monotonic operation. Focussing on (b) then, let us define $g := \bigwedge_i f_i$. Take a chain $x \leq x' \leq x''$. We have: $g[x'] = \bigwedge_i f_i[x'] \geq \bigwedge_i \left(f_i[x] \wedge f_i[x'']\right)$ (because each f_i is connected). Therefore $g[x'] \geq \bigwedge_{i,j} \left(f_i[x] \wedge f_j[x'']\right)$ (a conjunction with more terms is smaller). Therefore $g[x'] \geq \left(\bigwedge_i f_i[x]\right) \wedge \left(\bigwedge_i f_i[x'']\right)$. And so $g[x'] \geq g[x] \wedge g[x'']$. □

(57) **Result:** Disjunctive sentences

a. If f_i is a collection of monotonic functions, all increasing or all decreasing, $\bigvee_i f_i$ is monotonic in the same direction.
b. If ϕ and ψ are two elements that are both connected, $\phi \vee \psi$ isn't necessarily connected.

(58) **Corollary:** Existentially quantified sentences

[18] Note that this complicated condition is equivalent to "$\tilde{\phi}$ is monotonic", where: $\tilde{\phi} = \lambda w \cdot \{e_0 / [\![\phi[y \leftarrow e_0]]\!](w)\}$ (this is an object of type $\langle s, \langle e, t\rangle\rangle$). The same is true if "monotonic" is replaced by "connected".

a. If ϕ has a free variable y, and if for all individuals e_0, $\phi[y \leftarrow e_0]$ is monotonic in the same direction, then $\exists y, \phi$ is monotonic.
b. If ϕ has a free variable y, and if for all individuals e_0, $\phi[y \leftarrow e_0]$ is connected, then $\exists y, \phi$ isn't necessarily connected.

Proof. [Proof of (57)] Result (a) is an immediate consequence of the fact that disjunction is an increasing monotonic operation.

(b) Choose three elements x_1, x_2, x_3 such that $x_1 < x_2 < x_3$, and define the following functions over that input set and onto truth-values (so these are sets): $\phi = \{x_1\}$ and $\psi = \{x_3\}$. Both are monotonic (albeit in different directions), and therefore connected. Yet, $\phi \vee \psi = \{x_1, x_3\}$ is not connected. □

(59) **Result:** Exhaustified sentences

a. If ϕ and all of its alternatives are monotonic in the same direction, then EXHϕ is connected, but not necessarily monotonic.
b. If ϕ and all of its alternatives are connected, EXHϕ isn't necessarily connected.

Proof. For (a), assume without loss of generality that ϕ and its alternatives are upward monotonic. For any $\phi' \in$ ALT(ϕ), $\neg\phi'$ is decreasing monotonic, and therefore connected. Because conjunction preserves connectedness, it follows that EXHϕ is connected. The following counter example shows that EXHϕ, however, may not be monotonic when ϕ is: assume that ϕ's domain is $\{x_1, x_2, x_3\}$ such that $x_1 < x_2 < x_3$. Take $\phi = \{x_2, x_3\}$, $\phi' = \{x_3\}$, and ALT(ψ) = $\{\phi'\}$. Then EXH$\phi = \{x_2\}$, which isn't monotonic.

For (b), we will use the following counter-example, take: $\phi = \{x_1, x_2, x_3\}$, $\phi' = \{x_2\}$, and ALT(ϕ) = $\{\phi'\}$. Then EXH$\phi = \{x_1, x_3\}$ which is not connected. □

3.4.4 The CMH and the distribution of EXH

Thus, we may relate the theorems of Section 3.4.3 to generalizations discussed in Sections 3.1 and 3.2, and summarized in Table 3.1.

To do so, first suppose that primitive predicates are monotonic.[19] Then note that EXH transforms monotonic constituents into just connected constituents (see Result (59a)). Upon some further embedding then, the outcome may be non-connected at all, namely for those embeddings with a negative result about the preservation of connectedness (the b-part of the results in Section 3.4.3). This is the case for negation ((54b)), disjunction ((57b)) or existential contexts ((58b)). In other words, EXH in these environments may produce non-connectedness and therefore violate the CMH. In conjunctions or in the scope of universals however, connectedness is preserved ((55b), (56b)), and this guarantees that EXH will not by itself create non-connectedness, that is, it should comply with the CMH. Although these concern results about our idealized fragment, they correspond neatly with the corresponding generalizations from Table 3.1 about negation ((8)), disjunctions ((9), (10)), and existential environments ((11)).

Certain other examples in Table 3.1 are not immediately covered by this fragment. For instance, we did not include non-monotonic quantifiers, as in (6a) and (6b). This is essentially for the sake of brevity, as they could be treated in a similar fashion. Doing so would lead to the conclusion that EXH may lead to non-connectedness when embedded in a non-monotonic environment (see (6a)). In other words, non-monotonic environments should pattern just like negation or existential quantification, which is certainly among the less satisfying predictions of the CMH.

[19] In particular, if we order situations based on what set of propositions from a certain class they make true, and if among the propositions that we consider, there are all statements of the kind $P(x)$ for an individual x, then the predicate P will always be monotonic. This is always the case in examples that we consider.

The case of Free Choice inferences, in its traditional EXH treatment, also escapes from the fragment above. To fully treat these cases, the semantics of EXH would have to be made more precise and sophisticated (cf. for instance Fox 2007, and discussion in Spector 2016). That being said, let us simply accept that the resulting reading with a free choice inference is monotonic ($\Diamond A \wedge \Diamond B$ is monotonic, certainly if we consider the modal $\Diamond$ as an existential quantifier). Then the reason why it can be embedded under negation without creating non-connectedness is that it is a case where EXHϕ is not only connected, but also monotonic. Connected sentences may yield non-connected meanings under negation (Result (54b)), but monotonic ones may not (Result (54a)).

Hence, the current paradigm explains why certain occurrences of EXH operators can appear under negation, and others may not, independently of the fact that the resulting reading is weaker than the reading obtained without EXH. Overall, the generalizations from Table 3.1 correspond to the formal properties of connectedness, with the sole exception of the case of non-monotonic quantifiers.

3.5 Conclusions

The need for a restriction on the distribution of embedded exhaustification is a well-known problem of the grammatical theory of implicatures. We have formalized the Connectedness Meaning Hypothesis. This principle relies on the premise that non-connectedness create interpretive difficulties. The principle then proposes that parses with EXH which break connectedness are deviant. This proposal differs from previous approaches (referred to as the SMH or the EC) in two ways: first, these previous approaches rely on monotonicity and logical strengths considerations and, second, they rule out a given parse based on how it compares to another potential parse.

On the empirical side, the current proposal captures essentially the same generalizations as its predecessors, which is a good feature. We exhibited however a case where the proposals differ: previous accounts cannot tolerate the presence of EXH under negation, while the CMH allows it when EXH creates free choice inferences, and this seems to be an accurate result (see also Solt and Waldon (2019) for an application of similar ideas to interpretation patterns of sentences involving bare numerals).

Constraints based on monotonicity and logical strength are natural in formal semantics, and the move to connectedness considerations may sound unwarranted, if not suspicious. It is therefore worth pointing out that connectedness can be seen as a weak form of monotonicity. Importantly, in other domains, connectedness-based filters had been proposed as natural, cognitive tendencies, and they have been argued to be capable of explaining restrictions in the lexicons of content words (Gärdenfors 2004) and more recently in the lexicons of logical words (Chemla, Buccola, and Dautriche 2019). Connectedness filters are therefore natural constraints to explore, and we show here how they could have applications at the level of plain propositional meanings.

3.A Detailed treatment of some of the examples

In this section we go over some of the examples of Section 3.2.3 in more detail, showing why certain meanings are or aren't connected under the definition of connectedness in Section 3.4.2.

3.A.1 Example: basic negation

We begin by the basic example (60), repeated from (22).

(60) Sentence: It's not true that John did some of his homework.
Unattested parse: *It's not true that John did EXH [some of his homework].
Intended meaning: It isn't true that John did some but not all of his homework.

Making the simplifying assumptions that John's homework consists in two exercises, that John never does part of an exercise without finishing it, and that there are no other relevant things John could have done (or that none of these are crucial), we suppose that the appropriate order on situations is the one in Figure 3.2, where the situations are as in (61). This could be derived in the way described in Section 3.4.2 if one takes the QUD to be something like "What did John do?".

(61) $w_\varnothing$: John did neither exercise 1 nor exercise 2.
w_1 : John did exercise 1, but not exercise 2.
w_2 : John did exercise 2, but not exercise 1.
$w_{1,2}$: John did both exercise 1 and exercise 2.

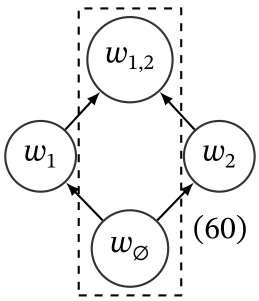

Figure 3.2 – Representation of the order over the situations in (61) for a sentence such as (60). The meaning of (60) is represented with the dotted box (it contains all situations that make the sentence true).

Figure 3.2 also indicates as a box the meaning of (60): one can see that this meaning is non-connected because the chain $w_\varnothing \to w_1 \to w_{1,2}$ has its extremities in the set and its in-between element out of it (see also Figure 3.1). On the other hand, the meaning one would obtain by parsing the sentence without EXH only contains $w_\varnothing$ and is therefore trivially connected.

3.A.2 Example: double disjunction

We now move on to (62), repeated here from (24). We take the order on worlds to be the one in Figure 3.3, which could derive from a question like "Who will John meet?". The labels of the situations correspond to the set of people that John meets. For the sake of simplicity we ignore people other than Ann, Bill and Carol (which does not affect the result).

(62) John will meet Ann, Bill or Carol.

a. Attested parse: EXH$[A \lor B \lor C]$
Meaning: John will meet exactly one person within Ann, Bill and Carol.

b. Unattested parse: *EXH$[A \vee$ EXH$[B \vee C]]$.
Intended meaning: John will meet exactly one person within Ann, Bill and Carol, or all of them.

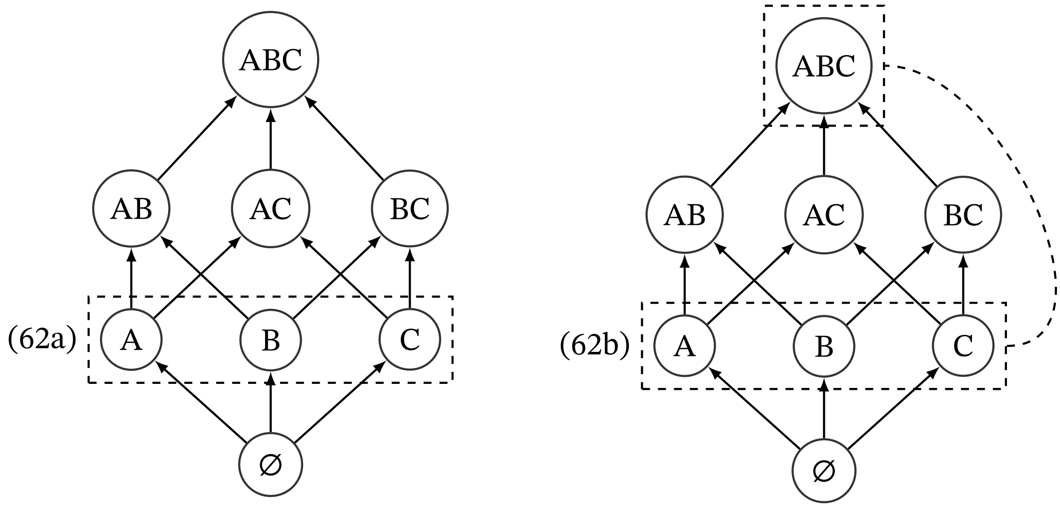

Figure 3.3 – Order on situations for (62) and meanings of parses of interest.

It can be seen from Figure 3.3 that (62a) is connected, as there is no non-trivial chain with its extremities in the truth box. On the other hand, for (62b), a number of chains have extremal points in the truth box and have in-between points outside of it ($A \rightarrow AB \rightarrow ABC$ is an example). Thus (62b) is not connected.

3.A.3 Example: existential sentences

As a third example, we consider (63), repeated from (26a). Recall that it could have all the interpretations in (64). Assuming that there are 2 employees A and B and 2 criminals 1 and 2, we take it that the appropriate order on situations is that of Figure 3.4, which could for instance derive from a QUD like "Which criminals did each employee see?". The labels of situations correspond to the extension of "see" (for example 'A1,A2,B1' corresponds to a situation in which Employee A saw Criminal 1 and Criminal 2, while Employee B only saw Criminal 2). One may check on Figure 3.4 that (64c) is non-connected, while (64b) is (the other two cases are omitted for the sake of clarity).

(63) a. There are employees who saw some of the criminals.
b. $\exists e, \exists c, S(e, c)$

(64) a. * $\exists_1 e$, EXH$_2 \exists_2 c, S(e, c)$
Some employees saw some but not all criminals.
b. EXH$_{1,2} \exists_1 e, \exists_2 c, S(e, c)$
Some but not all employees saw some criminals, and no employee saw all criminals.
c. * EXH$_1 \exists_1 e$, EXH$_2 \exists_2 c, S(e, c)$
Some but not all employees saw some but not all criminals.
d. EXH$_2 \exists_1 e, \exists_2 c, S(e, c)$
Some employees saw some criminals, but no employee saw all criminals.

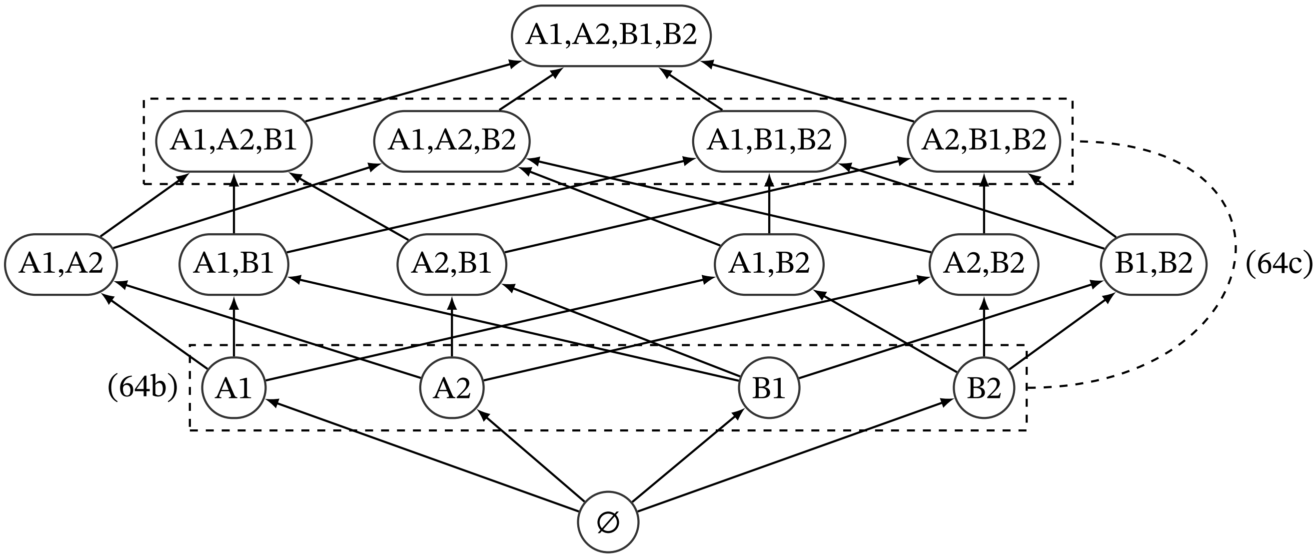

Figure 3.4 – Order on situations and meanings for (63).

3.A.4 Other examples

The cases of universal quantification and non-monotonic quantification ((29a) and (31)) are not detailed here for the sake of brevity. We simply just note that we assume the relevant order for them is something similar to the one for (63).

3.B Other notions of in-betweenness and connectedness

In this section, we discuss three alternatives to our definitions of in-betweenness and connectedness, and show why they seem inappropriate for the purpose at hand. These alternatives seem to be the most natural other choices we could have made, but we make no claim of comprehensiveness.

3.B.1 In-betweenness based on minimal paths

A natural definition of in-betweenness in metric spaces is the following: x is in-between y and z if it is on a minimal-distance path from y to z.[20] This requires a notion of distance. How can this apply to situations? Let us take the example of the four situations of (60):

(65) $w_\varnothing$: John did neither exercise 1 nor exercise 2.
w_1 : John did exercise 1, but not exercise 2.
w_2 : John did exercise 2, but not exercise 1.
$w_{1,2}$: John did both exercise 1 and exercise 2.

A natural notion of distance can be inferred from a diagram like Figure 3.5, where the distance between two nodes is given by the minimal number of edges that must be traversed to go from one to the other. Thus, w_1 and w_2 are 2 units apart; both paths joining them are minimal, so that both $w_\varnothing$ and $w_{1,2}$ are in-between them. As a result, even the most basic sentence, "John did some of his homework", is non-connected with or without its

[20]With this definition of in-betweenness, what we call connectedness corresponds to what is usually called convexity in mathematics.

exhaustified meaning, because it is true in both w_1 and w_2, but not in $w_\varnothing$. Hence, this notion of connectedness surely is too stringent for our current purposes.[21]

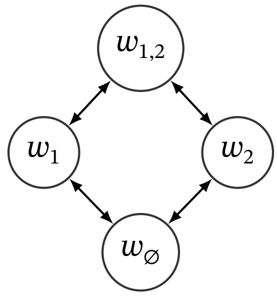

Figure 3.5 – A possible way to map the possible worlds in (65) into a feature space with a notion of distance.

3.B.2 Standard notion of connectedness on undirected graphs

An alternative approach also not based on an order, would be to take the same graph structure as from 3.B.1, and define connectedness as such: if there are two worlds where the proposition is true, there must be *some* path from one to the other throughout which the proposition stays true. Here the notion of in-betweenness does not play a direct role. We note that this is closer to some usual definitions of connectedness in topology.

This notion is less stringent than the one above in 3.B.1. Under this view, the non-exhaustified (and in fact monotonic) reading of "John did some of his homework" would be connected, but its exhaustified meaning, again, would not be. In fact, Result (59) would not hold, and none of the usual logical operations (conjunction, disjunction, etc.) preserve this notion of connectedness. Hence, again, this notion cannot plausibly constrain the distribution of EXH.

3.B.3 Weak order-based connectedness

We may adopt our order-based definition of in-betweenness, but also a weaker notion of connectedness: instead of requiring that every world in-between two true worlds make the proposition true, we could require that *at least some* of them do. This notion is strictly weaker than ours. It has some merits: it doesn't change anything for basic examples such as (60), and it makes Result (59) (about how EXH preserves monotonicity and connectedness) true. We haven't fully investigated whether this notion's empirical predictions would be better or worse than those of the definition we adopted; in general, as it is a weaker constraint, it ought to ban fewer parses, and our cursory look suggests it is too lax: for instance, it doesn't ban (66).

(66) Parse: *John didn't *exh* [do exercise 2.]
Intended meaning: it isn't the case that exercise 2 is the exact extent of what John did.

[21] Adding an edge between w_1 and w_2 may solve the problem, but other issues would still arise in more complicated cases. For instance, one may verify that (27d) would not be connected in this sense. More generally, this notion does not seem to make a result like (59) true; i.e., it is not guaranteed that global exhaustification of monotonic components will be connected.

References

Barwise, Jon and Robin Cooper (1981). “Generalized quantifiers and natural language”. *Linguistics and Philosophy* 4.2, pp. 159–259.

Beaver, David (2001). *Presupposition and Assertion in Dynamic Semantics*. CSLI Publications.

Chemla, Emmanuel (2009). “Universal Implicatures and Free Choice Effects: Experimental Data”. *Semantics and Pragmatics* 2.2, pp. 1–33.

Chemla, Emmanuel, Brian Buccola, and Isabelle Dautriche (2019). “Connecting Content and Logical Words”. *Journal of Semantics* 3, pp. 531–547.

Chemla, Emmanuel and Benjamin Spector (2011). “Experimental evidence for embedded scalar implicatures”. *Journal of Semantics* 28.3, pp. 359–400.

Chemla, Emmanuel et al. (2018). “Constraints on the lexicons of human languages have cognitive roots present in baboons (*Papio papio*)”. Ms. CNRS, University of Edinburgh, Aix-Marseille University.

Chierchia, Gennaro, Danny Fox, and Benjamin Spector (2011). “The grammatical view of scalar implicatures and the relationship between semantics and pragmatics”. In: *Semantics: an international handbook of natural language meaning*. Ed. by Paul Portner, Claudia Maienborn, and Klaus von Heusinger. Vol. 3. De Gruyter Mouton, pp. 2297–2332.

Coppock, Elizabeth and David Beaver (2014). “Principles of the Exclusive Muddle”. *Journal of Semantics* 31.3, pp. 371–432.

Dalrymple, Mary et al. (1998). “Reciprocal expressions and the concept of reciprocity”. *Linguistics and Philosophy* 21.2, pp. 159–210.

von Fintel, Kai and Lisa Matthewson (2008). “Universals in semantics”. *The linguistic review* 25.1-2, pp. 139–201.

Fox, Danny (2007). “Free Choice Disjunction and the Theory of Scalar Implicatures”. In: *Presupposition and Implicature in Compositional Semantics*. Ed. by Uli Sauerland and Penka Stateva. Palgrave Macmillan, pp. 71–120.

Fox, Danny and Benjamin Spector (2018). “Economy and embedded exhaustification”. *Natural Language Semantics* 26.1, pp. 1–50.

Gärdenfors, Peter (2004). *Conceptual spaces. The geometry of thought*. MIT press.

Gärdenfors, Peter (2014). *The geometry of meaning: Semantics based on conceptual spaces*. MIT Press.

Gotzner, Nicole and Anton Benz (2018). “The Best Response Paradigm: A New Approach to Test Implicatures of Complex Sentences”. *Frontiers in Communication* 2 (21).

Heim, Irene (1983). “On the projection problem for presuppositions”. In: *Formal Semantics: The Essential Readings*. Ed. by Paul Portner and Barbara Partee. Blackwell, Oxford, pp. 249–260.

Horn, Laurence R. (1973). “On the semantic properties of logical operators in English”. PhD thesis. University of California.

Katzir, Roni (2007). “Structurally-defined Alternatives”. *Linguistic and Philosophy* 30.6, pp. 669–690.

Katzir, Roni and Raj Singh (2013a). “Constraints on the lexicalization of logical operators”. *Linguistics and Philosophy* 36.1, pp. 1–29.

Katzir, Roni and Raj Singh (2013b). “Hurford disjunctions: embedded exhaustification and structural economy”. *Proceedings of Sinn und Bedeutung* 18, pp. 201–216.

Kratzer, Angelika (1979). “Conditional necessity and possibility”. In: *Semantics from Different Points of View*. Ed. by Rainer Baüerle, Urs Egli, and Armin von Stechow. Springer, Berlin, pp. 117–147.

Kratzer, Angelika (1981). “Partition and Revision: The Semantics of Counterfactuals”. *Journal of Philosophical Logic* 2, pp. 201–216.

Kratzer, Angelika and Junko Shimoyama (2002). “Indeterminate pronouns: The view from Japanese”. In: *Proceedings of the Tokyo conference on psycholinguistics*. Ed. by Yukio Otsu. Vol. 3. Hituzi Syobo, pp. 1–25.

Krifka, Manfred (1993). “Focus and Presupposition in Dynamic Interpretation”. *Journal of Semantics* 10.4, pp. 269–300.

Lewis, David (1973). *Counterfactuals*. Basil Blackwell.

Lewis, David (1981). “Ordering semantics and premise semantics for counterfactuals”. *Journal of Philosophical Logic* 10.2, pp. 217–234.

Mayr, Clemens and Jacopo Romoli (2016). “A puzzle for theories of redundancy: Exhaustification, incrementality, and the notion of local context”. *Semantics and Pragmatics* 9.7, pp. 1–48.

Potts, Christopher et al. (2016). “Embedded Implicatures as Pragmatic Inferences under Compositional Lexical Uncertainty”. *Journal of Semantics* 33.4, pp. 755–802.

Roberts, Craige (1996). “Information structure in discourse: towards an integrated formal theory of pragmatics”. *OSU working papers in linguistics* 49, pp. 91–136.

Roberts, Craige (2006). “Only, presupposition and implicature”. Ms., The Ohio State University.

van Rooij, Robert and Katrin Schulz (2004). “Exhaustive interpretation of complex sentences”. *Journal of logic, language and information* 13.4, pp. 491–519.

Sauerland, Uli (2012). “The computation of scalar implicatures: Pragmatic, lexical or grammatical?” *Language and Linguistics Compass* 6.1, pp. 36–49.

Schlenker, Philippe (2008). “Be articulate: A pragmatic theory of presupposition projection”. *Theoretical Linguistics* 34.3, pp. 157–212.

Schulz, Katrin and Robert van Rooij (2006). “Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation”. *Linguistics and Philosophy* 29.2, p. 205.

Schwarzschild, Roger (1999). “GIVENness, AvoidF and other constraints on the placement of accent”. *Natural language semantics* 7.2, pp. 141–177.

Solt, Stephanie and Brandon Waldon (2019). “Numerals under negation: empirical findings”. *Glossa: A Journal of General Linguistics* 4 (1), p. 113.

Spector, Benjamin (2006). “Aspects de la pragmatique des opérateurs logiques”. PhD thesis. Université Paris 7.

Spector, Benjamin (2016). “Comparing exhaustivity operators”. *Semantics and Pragmatics* 9.11, pp. 1–33.

Stalnaker, Robert (1968). “A Theory of Conditionals”. In: *IFS*. Ed. by W. Harper, R. Stalnaker, and G. Pearce. Springer, Dordrecht, pp. 41–55.

Chapter 4

Minimal sufficiency readings of necessity modals

This chapter is an as yet unpublished manuscript. I presented related work at NELS 50, and an abridged version of the chapter was published as:

> Émile Enguehard (2020). "Minimal sufficiency readings of necessity modals". In: *Proceedings of NELS 50*. Ed. by Mariam Asatryan, Yixiao Song, and Ayana Whitmal. Vol. 1, pp. 217–226.

Abstract

This chapter discusses a series of cases where a teleological necessity modal appears to receive a weaker meaning than expected, including the "Sufficiency Modal Construction" of von Fintel and Iatridou (2007). I show that these examples express a shared meaning, Minimal Sufficiency, and that describing this meaning requires reference to a ranked scale. The chapter discusses three analytical approaches: von Fintel and Iatridou's decompositional analysis of *only*, an account based on a scalar weakening operator in the spirit of Alonso-Ovalle and Hirsch (2018), and an account based on alternative-sensitive modals in the spirit of Krasikova (2010). None of the approaches is found to be fully satisfying. The discussion highlights the various challenges raised by Minimal Sufficiency examples.

Acknowledgements

I thank Benjamin Spector, Jakub Dotlačil, Brian Buccola, Sabine Iatridou, Kai von Fintel, Roger Schwarzchild, and the audiences at PALLMYR XII and NELS 50 for discussion.

4.1 Introduction: issues with purpose constructions

4.1.1 The prejacent problem: where to find cheese in Boston

The combination of *have to* with an infinitive purpose clauses, as in (1), has been called teleological (or goal-oriented) *have-to*.

(1) To get a driving license, you have to be an adult.

A possible analysis of (1) is one where the infinitive clause introduces an ordering on possible worlds, and the modal is a universal quantifier (cf. for instance von Fintel and Iatridou 2005). Then, (1) can be paraphrased as (2).

(2) In all the best worlds w.r.t. getting a driving license, you are an adult.[1]

However,von Fintel and Iatridou (2007) point out that this paraphrase seems incorrect for (3a). According to the analysis of *only* of Horn (1969), (3a) ought to presuppose the truth of (3b). (3b) would in turn have a paraphrase similar to (2).

(3) a. To get good cheese, you only have to go to the North End [a part of Boston].
b. To get good cheese, you have to go to the North End.

However, if we know that there are many places in Boston where you can get good cheese, we can accept (3a) but not (3b). The falsity or infelicity of (3b) is expected in such a context given the paraphrase in (2), but (3a) should be rejected as well. This is called the *prejacent problem* by von Fintel and Iatridou (2007) (as (3b) is the *prejacent* of (3a)). The usage of *have to* that (3a) exemplifies is dubbed the "sufficiency modal construction" (SMC) by von Fintel and Iatridou (2007). Given that the difference between (3a) and (3b) lies in the presence of *only*, the analysis that they propose relies chiefly on a novel analysis of *only*.

In this article, we will relate the SMC to another family of examples involving teleological modals, but not necessarily *only*, and we will call the resulting class of examples *Minimal Sufficiency statements*. What brings these modal examples together is that they license a certain set of inferences about propositions lying on a scale, which inferences we will call *Minimal Sufficiency* (MS).

In general, MS statements offer a challenge to existing views on modality, as well as operators like *only*. Indeed, the MS inference is different to what our established understanding predicts, and MS statements are judged to be acceptable in contexts where one would predict them to be false or presuppositional failures. Accounting for this fact while deriving the MS inference in full poses interesting challenges, particularly as concerns the division of labour between semantics and pragmatics. In a sense, the identification of MS statements can be seen as a generalization of the prejacent problem.

We are going to explore here several natural routes towards an account of MS statements. All of them involve a significant weakening of a certain element of the construction, relative to established views, albeit a different one. We will see that all of our attempts encounter serious issues, highlighting the continued difficulty of the problem at hand.

[1] In this paraphrase, we implicitly adopt the *Plural Limit Assumption*: the ordering of worlds is such that there is a set of *best worlds*, i.e. minimal elements for the order, and that all worlds are ranked above at least one of the best worlds. As far as we can tell, our discussion does not hinge on the Limit Assumption.

4.1.2 Some new examples, and an outline

The two examples that follow are very similar to (3a): they also appear to defeat the paraphrase of modality that we started with, in that whatever constraints they set on the world are much weaker than what one would predict. In (4), it is not implied that all situations where you get into the next stage are situations where you get a silver medal. Specifically, (4) is compatible with there being possible situations where you get a gold medal and get into the next stage (in fact (4) licenses a stronger inference: all situations where you get a gold medal are such that you get into the next stage).

(4) To get into the next stage, you have to get a silver medal.

In (5) (where capital letters indicate prosodic prominence), it again fails to be implied that you cannot be over the legal limit without drinking two beers. For instance, you may probably be over the legal limit after having drunk a certain amount of wine and no beer at all. (5) doesn't intuitively contradict this fact, even though we would predict it to do so. This failure of necessity is very reminiscent of (3a) and (4) above.

(5) To be over the legal limit, you have to drink TWO beers.

The similarities in meaning between (3a), (4) and (5) lead us to categorize them together as expressing Minimal Sufficiency (MS). In the next section, we will detail what Minimal Sufficiency is and what is challenging about it. We will then review the analysis of von Fintel and Iatridou (2007) for (3a), consisting in assuming that *only* can be decomposed, and show that it does not provide an adequate characterization of what MS means, although part of the problem can be resolved with some simple modifications to the analysis. Additionally, since this approach relies primarily on a novel account of *only*, it requires additional assumptions to extend to (4) and (5), where *only* is not present.

If the explanation for failures of necessity is not to be found in a weakening of *only*, there are two other options: weakening the embedded clause, and weakening the modal. We are going to explore both options in that order, and show that while the embedded-clause option faces some challenges, the modal option is much more difficult to make viable than one would assume, perhaps pointing towards the embedded-clause solution.

In Section 4.3, we will start with discussing an analysis of the cheese example (3a) proposed by Alonso-Ovalle and Hirsch (2018), consisting in assuming that the embedded clause is the scope of a covert weakening operator AT-LEAST. We will see why it needs to be tweaked to properly account for the full range of MS examples, by giving the operator degree-based semantics and making it even weaker than Alonso-Ovalle and Hirsch (2018) assume. We will mention some potential shortcomings of the resulting picture. The analysis discussed in Section 4.3 is an instance of the second route, where the embedded clause is weakened while the modal is kept strong.

Section 4.4 explores the third route: weakening the modal. In what we call an alternative-sensitive analysis, AT-LEAST is kept as Alonso-Ovalle and Hirsch (2018) define it, and we give alternative-sensitive semantics to modality, following Villalta (2008). While such an alternative-sensitive approach is attractive in principle, we will see that it runs into a number of challenges and we will not be able to provide a satisfactory analysis in this vein.

Finally, Section 4.5 discusses certain empirical aspects of MS constructions in more detail, in particular relating to the exact status of certain inferences when *only* is present in MS examples. We will show that our observations are a challenge for the various analyses that we discuss, and that they might help us gain more insight into the contentious issue of how to analyse *only*.

4.2 Minimal Sufficiency

4.2.1 A breakdown of Minimal Sufficiency

The reason we categorize (3a), (4) and (5) together is that their meanings have a number of shared components. First, they involve a scale, a set of ranked alternatives. In (4), the scale is the conventional scale of medals shown in (6). In (5), it is the scale of numerals (one, two, three…). In (3a), there is no obvious scale, but we are going to see that we cannot describe its meaning without assuming one is present. That scale features places one might go to to buy things and ranks them by the effort involved; for the sake of exposition we can assume that it looks like (7a), and is ranked as in (7b).

(6) bronze medal $\prec$ silver medal $\prec$ gold medal

(7) a. $A = \left\{ \begin{array}{l} \phi_C \,:\, \text{you cross the street,} \\ \phi_{NE} \,:\, \text{you go to the North End,} \\ \phi_{SB} \,:\, \text{you go to South Boston} \end{array} \right\}$,

b. $\phi_C \prec \phi_{NE} \prec \phi_{SB}$.

Given a scale, MS consists in the conjunction of *Sufficiency* and *Minimality*. Sufficiency is the inference that the alternative being considered is a way to attain the goal, while higher alternatives are unnecessary. (4) implies in some way that getting silver will in fact let you qualify, while getting gold is not necessary. Similarly, (5) implies that drinking two beers is enough to put you over the limit. (3a) also implies that you can get good cheese in the North End.

Minimality is the inference that lower alternatives are not a way to attain the goal. (5) implies that drinking one beer will leave you below the limit, and (4) implies that getting bronze will not let you qualify. In (3a), it is less obvious that Minimality is present, but it can still be detected. Assume that you can get good cheese in many places in Boston; the North End is one, and another one is the cheese shop just across the street from where the speaker resides; the speaker is fully informed about this situation. In a situation like this, it seems that it would be odd for the speaker to say (3a). Intuitively, the element in this situation that is incompatible with (3a) is that there is another way of getting good cheese that is clearly easier than going to the North End. Thus, while (3a) is compatible with there being good cheese shops outside of the North End, it suggests in some way that the North End is amongst the most accessible places where one can get good cheese, or in other words that other places where you can get good cheese are at least as difficult to access as the North End. This is the meaning component that we call Minimality.

It is important to note that our assumption that (3a) involves a scale is essential to even just describe Minimality. In our description and in our analysis, we need to make reference to a scale to account for the difference between "you can find good cheese in South Boston" (whose truth is perfectly compatible with that of (3a)) and "you can find good cheese across the street" (whose truth seems to make (3a) deviant).[2]

[2] The status of Minimality in (3a) is in fact not clear: is it an assertion, a presupposition, a scalar implicature? We will discuss the issue in Section 4.5. However, nothing we say here depends on what that status is: our point is that Minimality exists in some form, which is seen in the fact that (3a) is deviant *in some way* if you can get good cheese across the street, and that scalarity is needed to derive Minimality *in any form*.

The meaning of our examples does not appear to be any stronger than MS as we just described it. In particular, there is nothing resembling *necessity* as we paraphrase it in (2). The examples fail to implicate necessity of the condition they set in two ways: first, they allow for higher alternatives (going to South Boston, getting gold, drinking exactly three beers[3]) to also let you reach the goal, and in fact implicate that they do. Second, they allow for alternatives not on the scale at all to let you reach the goal. (3a) does not imply that you cannot find good cheese without going anywhere: for instance, you might perhaps ask your neighbour for some. (4) is compatible with a situation where, if you fail to get the silver medal, you may still get into the next stage through some unrelated means, such as being invited. Most clearly, (5) is compatible with the fact that you may drink only wine and end up above the limit.

These commonalities group our three examples (3a), (4) and (5) together as instances of the MS reading of necessity modals. In contrast, (1) does not involve a scale, does not mean that you automatically get awarded a driving license when you turn 18 (no sufficiency), and does mean that if you're not an adult you cannot get a driving license (necessity).

4.2.2 The decompositional analysis

Before we move on to our proposals, we must discuss that of von Fintel and Iatridou (2007). The difference between (3a) and (3b) being the presence of *only*, von Fintel and Iatridou (2007) propose that the prejacent problem can be solved by reanalysing *only*. They take it to be decomposable into two elements, a negation NEG and an exceptive QUE paraphrasable as "anything other than".

(8)
$$\text{only} = \text{NEG} + \text{QUE}$$
$$[\![\text{NEG}]\!] = \lambda w.\lambda p.\,[p(w) = 0]$$
$$[\![\text{QUE}\ p]\!] = \lambda w : [\exists p' \in \text{ALT}(p),\ p'(w) = 1].\,[\exists p' \in \text{ALT}(p),\ p' \neq p \wedge p'(w) = 1]$$

In these definitions, $\text{ALT}(p)$ denotes the set of formal alternatives to p, in the sense of alternative semantics (Rooth 1992). The decompositional view of *only* opens up the possibility that in (3a), there is a split-scope configuration:

(9) NEG $\gg$ have-to $\gg$ QUE

Von Fintel and Iatridou (2007) assume that the split-scope configuration is in fact realized in (3a) under the relevant reading. They further assume, following Horn (1996), that the presupposition of *only* is a mere existential presupposition (some alternative to the prejacent is true), and also that it is in fact que that carries this presupposition. Then, (3a) presupposes that in all the best worlds where you get good cheese, you go somewhere, which is almost trivial, and asserts that in not all of these worlds you go somewhere that is not the North End, i.e. that you can go to the North End and get good cheese.

The analysis just presented predicts very weak truth and usage conditions for (3a). In particular, it does not derive Minimality. Recall that we called Minimality the fact that (3a) is deviant in a situation where you can in fact get good cheese across the street. In the decompositional analysis, however, (3a) would be a perfectly felicitous and true utterance in

[3]Note, though, that as long as we adopt an "at-least" semantics for bare numerals (cf. Spector 2013 for discussion of the issue and arguments in this direction), "drinking exactly three beers" is a subset of "drinking two beers" and the fact that you can drink three beers is not a violation of necessity for (5). We list this case anyway to highlight the parallels between the examples.

such a context. Recall also that it is impossible to derive Minimality without making reference to a scale, because we need to find a difference between "you can go across the street and get good cheese" and "you can go to South Boston and get good cheese": only the latter is compatible with the truth of (3a). As there is no scalar component in the decompositional analysis, there is no way to derive that fact.

To be sure, von Fintel and Iatridou (2007, sec. 4.2) remark that in general, "[*only* is] associated with a scale and the focus is low on the relevant scale". While they do not elaborate on how this fact relates to their non-scalar semantics, they propose that it explains why (3a) suggests that going to the North End is easy; let us call this component of the meaning of (3a) "Easiness." Von Fintel and Iatridou (2007) provide empirical as well as theoretical arguments (based on compositionality) to the effect that what Easiness means is that going to the North End is easy relative to actions in general, and not relative to ways to get good cheese. If that is so, Easiness is perfectly compatible with violations of Minimality: we may very well imagine that going to the North End is an easy thing to do, for instance because the speaker lives a couple of subway stops away from the North End, and yet even in that context, (3a) implies that you cannot find good cheese across the street. Thus, we cannot derive Minimality from Easiness.[4]

Von Fintel and Iatridou (2007, sec. 4.3) also briefly sketch a scalar version of their proposal, and note that it might help explain why "[sufficiency modals] seem to rate the ways of achieving the goal and zero in on the easiest, least-effort-involving way" (this is Minimality). They reject it for unrelated reasons. Since we think that deriving Minimality is important enough that one should pursue the effort, we will attempt here to "fix" the decompositional analysis and make it scalar.

Let us assume, then, that "go to the North End" lies on a pragmatic scale like the one in (7a) for the purposes of interpreting (3a). We can look at existing proposals on the interaction between *only* and pragmatic scales, which has been the subject of a great amount of literature (see in particular Klinedinst 2004). Of particular interest to us is Greenberg's (2019) proposal that *only* presupposes that the prejacent is the lowest alternative being considered. Adopting this proposal within the decompositional analysis of *only* is most naturally done through the following entry for QUE:

(10) $$[\![\text{QUE } p]\!] = \lambda w : [\exists p' \in \text{ALT}(p),\ p'(w) = 1] \wedge [\forall p' \in \text{ALT}(p),\ p' \geq p].$$
$$[\exists p' \in \text{ALT}(p),\ p' \neq p \wedge p'(w) = 1]$$

We then predict that for (3a) to be felicitous, the alternatives to "go to the North End" must not include anything "easier". The existential presupposition of QUE, applied to these alternatives, will end up being that at least one of them is true. This presupposition projects universally from the scope of the modal, resulting in the overall presupposition for (3a) that in all worlds where you get good cheese, you go somewhere that is at least as inaccessible as the North End.[5] In the problematic situation where you can in fact get good cheese across the street, we predict (3a) to be a presupposition failure.[6] Thus, we do get Minimality.

[4] Our breakdown of MS did not include Easiness, as the reader will have noticed. This is because no trace of Easiness is detectable in MS examples that do not feature *only*, such as (4) and (5). Thus, we follow von Fintel and Iatridou (2007) in assuming that Easiness is contributed by *only* and in not analysing it further.

[5] In general, the combination of Greenberg's scalar presupposition (the prejacent is the lower end point), and an existential presupposition over the same set of alternatives, results overall in an "at least" presupposition for *only*. D. I. Beaver and Clark (2008) independently argue for such a presupposition.

[6] As we already noted, the status of the Minimality inference in (3a) is not clear. We will discuss it some more in Section 4.5. We can however leave this issue aside for now, as we are going to discuss alternatives to

What we have just seen is that the decompositional analysis can be "fixed" to derive MS for (3a). However, even when fixed, the decompositional account will not extend to examples not containing *only* such as (4) and (5). For this reason, we will explore different ways to account for MS in the next sections. While doing so, we will sometimes make reference to the "decompositional analysis", by which we mean the "fixed" scalar version.

4.3 Getting MS by weakening the condition

4.3.1 Pragmatic scales, weakening operators, and exhaustification

One of the assumptions made in the decompositional analysis is that *only*'s presupposition is weaker than proposed by Horn (1969). Because they find arguments against this assumption, Alonso-Ovalle and Hirsch (2018) propose a different analysis of the prejacent problem involving a weakening operator AT-LEAST (the introduction of which they attribute to Schwarz (2005)).

The idea behind AT-LEAST is that it acts upon propositions lying on a pragmatic scale, and returns their disjunction with any higher-ranked alternatives. A possible entry is given in (11). It omits the necessary presupposition that the alternatives should form a scale.

(11) $[\![\text{AT-LEAST } p]\!] = \lambda w.\, \exists q \in \text{ALT}(p),\ [q \geq p \wedge q(w) = 1]$.[7]

Alonso-Ovalle and Hirsch (2018) propose that (3a) has an LF similar to (12). Under this analysis, assuming the relevant scale is the one in (7b), what the prejacent asserts is that in all worlds where you get good cheese, you go to the North End or you do something more difficult, such as going to South Boston. It follows that crossing the street won't do (Minimality), but not that going to the North End is necessary (failure of Necessity). Because *only* presupposes the truth of its prejacent, all this is presupposed in (3a), which further asserts that replacing "the North End" by any higher-ranked alternative, and in particular by "South Boston" yields a false sentence, i.e. that it is not the case that in all worlds where you find good cheese you go to South Boston or some more difficult place. It follows, then, that in some worlds where you get good cheese, you go to the North End exactly (Sufficiency).

(12) To get good cheese, you only have to AT-LEAST [PRO go to the NE].

The fact that what looks like the prejacent of (3a), (3b), doesn't receive an MS reading leads us to conclude that it doesn't have an easily accessible reading where AT-LEAST is present. We can explain the contrast in the following way: AT-LEAST presupposes that its argument lies on a pragmatic scale, but "going to the North End" isn't a scalar item, making AT-LEAST unavailable in (3b). However, *only* presupposes that there is a contextual ranking available, making AT-LEAST licit.[8] Thus *only*'s only role in making (3a) get an MS interpre-

the decompositional analysis anyway.

[7] Both AT-LEAST and AT-LEAST-DEG, which we are going to introduce below, can be seen as instantiations of Beaver and Coppock's (2014) MIN operation; the crucial difference between the two will lie in the domain they quantify over.

[8] Another explanation in the same spirit, but based on a different treatment of *only* goes as follows: *only* does not in fact ever combine with pragmatic scales, but only with logical ones; it asserts that no logically stronger alternatives are true, and presupposes among other things that such alternatives exist. When the prejacent doesn't have obvious stronger alternatives, it is necessary to take it to be part of a scale and to insert AT-LEAST below *only* to satisfy this presupposition.

tation consists in forcing us to accommodate a ranking.[9]

This line of explanation leads us to the prediction that if a pragmatic scale is easily accommodated, *only* will not be necessary for AT-LEAST to occur. This prediction is borne out by (4). In (4), the ranking of alternatives is a well-established convention; this makes the insertion of AT-LEAST licit even without *only*, which is why an MS reading obtains. Thus the analysis extends to (4), which we take to have an LF similar to (13).[10]

(13) To get into the next stage, you have to AT-LEAST [PRO get a silver medal.]

One issue at this point is that our derivation of Sufficiency for (3a) involved the contribution of *only*. In (13), *only* is not present. The truth conditions we predict for (13) entail Minimality (getting bronze won't do), but they are compatible with Sufficiency being violated; concretely we predict (13) to be true if what you actually need is to get a gold medal. Intuitively, however, (13) is false in this case.

Fortunately for us, it is standard to assume that upon hearing (13), a hearer would conclude that the stronger sentence whose LF is given in (14) is false, as otherwise the speaker would have said it; in other words Sufficiency should follow from (13) as an implicature. One way to formalize the process by which implicatures are derived is to assume that all sentences can contain a silent counterpart of *only*, often denoted as EXH (after Groenendijk and Stokhof (1984)). Then the LF in (15) should deliver both Sufficiency and Minimality.[11] Since the Sufficiency inference appears to be obligatory here, we have to assume that EXH is obligatory in (4), a stipulation that we do not have an explanation for at this point.[12]

(14) To get into the next stage, you have to AT-LEAST [PRO get a gold medal.]

(15) ..., EXH [you have to AT-LEAST [PRO get a silver medal.]]

4.3.2 Failure of necessity and degrees

It seems that, at this point, we have a satisfactory analysis of MS with pragmatic scales involving a weakening of the condition through the weakening operator AT-LEAST. There is,

[9] As we already noted, *only* also contributes an "easiness" inference: going to the North End is easy. Since that inference is not found in the other examples, we take it to be an effect of *only* and not part of MS.

[10] Our explanation of the contrast between (3a) and (3b) also makes the prediction that the presence of other overt scalar operators than *only* should also license AT-LEAST. At first glance, this prediction is borne out. Consider (ia), where the overt scalar operator "at least" is present. Assuming that "at least" means the same thing as its covert counterpart, and that it takes scope above the modal, we predict (ia) to mean something like: to get good cheese, you have to go to the North End, or you have to go to South Boston, or you have to go to New York City, etc. In other words, there is a place at least as inaccessible as the North End such that you have to go there to get good cheese. This is of course not what (ia) means: (ia) does not imply that there is any one place that you absolutely have to go to to get good cheese. What it means is that the most easily reached place where one can find good cheese is at least as far as the North End. Though there might be other ways, we can derive the correct meaning if we just assume that the LF of (ia) is as in (ib). Thus it appears that in (ia) an overt "at least" above the modal licenses a covert AT-LEAST below it.

(i) a. To get good cheese, you at least have to go to the North End.
    b. ..., you at least have to AT-LEAST [PRO go to the North End.]

[11] Here we treat EXH as an element of the grammar, but nothing hinges on this at this point, and one may prefer to view it as a purely theoretical device standing in for a pragmatic process instead.

[12] Possibly, EXH is obligatory when there is no uncertainty as to what the alternatives are, and they form a non-trivial scale. Evidence in this direction can be found when looking at numerals, as we will discuss in Section 4.3.4.

however, a subtle issue. Recall that we divided "failure of necessity" into two components: failure of necessity w.r.t. scalemates and w.r.t. non-scalemates. The former is provided for by AT-LEAST: since AT-LEAST essentially replaces the condition with its disjunction with higher alternatives, the "have to" statement no longer rules out these alternatives. The latter is not dealt with. The problem is that MS statements also fail to rule out alternatives not found on the scale: (3a) is compatible with it being possible to find good cheese without going anywhere, perhaps if you steal your neighbour's, and (4) is compatible with it being possible to get into the next stage while having lost in the current stage, perhaps because you can be invited by the organizers.

In these two examples, the problematic alternatives can be argued to be remote possibilities, or atypical, or irrelevant, and it sounds like we should look for a solution involving these notions. This will however not work at all for (16) (repeated from (5)). (16) doesn't rule out being over the limit after having drunk some amount of wine, but drinking wine is not intuitively a remote or irrelevant possibility; in fact one may very well utter (16) while drinking wine.[13] Note also that, with an "at-least" semantics for numerals, failure of necessity w.r.t. higher scalemates (i.e. the lack of an inference that drinking three beers is out) is not puzzling in (16) and inserting AT-LEAST would be vacuous.

(16) To be over the legal limit, you have to drink TWO beers.

The problem, then, is that the weakening brought about by AT-LEAST is not sufficient for our purposes: the condition is still too strong. We want to make it so that the condition set in (16) doesn't rule out drinking wine; only lower scalemates, such as drinking exactly one beer, are to be ruled out.[14] A natural move is to also include the possibility of drinking wine in the grand disjunction that AT-LEAST represents, i.e., to make (16) paraphrasable as (17).

(17) ..., you have to drink two beers or a sufficient amount of something else.

However, recall that in the definition of AT-LEAST(p), we quantify over ALT(p), the set of formal alternatives to p. Following standard assumptions on the effect of prosodic focus (Rooth 1992), the prosody of (16) should constrain the formal alternatives to be as in (18).

(18) ALT("you drink TWO beers") = {"you drink n beers" | n}

Given that this prosody appears to be necessary for the MS reading to obtain (cf. discussion in Section 3.4), we cannot assume that the alternatives to "you drink two beers" in (16) include things like "you drink a glass of wine". Thus, we need to divorce AT-LEAST from the set of focus alternatives. We offer here an implementation of this move in terms of degrees.

Bale (2008) introduces the notion of a *universal degree scale*, an abstract scale (isomorphic to part or all of the real line) to which expressions of measure or evaluative expressions may refer. Universal degrees do not have to correspond to any physical or evaluative scale. This suggests that universal degrees may appear even in the semantics of expressions that are not obviously gradable.

Rett (2014) proposes that NPs may be be interpreted as denoting a degree (type d), which

[13] A potential solution involving some kind domain restriction naturally comes to mind, but there is reason to think it cannot be the end of the story; see Section 4.4.4 for discussion.

[14] The same problem would occur if we had adopted the fixed decompositional analysis, and extended it from *only* to EXH: we would predict (16) to presuppose that if you're over the limit, you have drunk two beers. The original decompositional analysis would correctly fail to derive this inference, but would also fail to derive Minimality in any of our examples.

allows them to be used in degree constructions like comparatives, and makes them trigger singular agreement; in (19) both phenomena can be seen.

(19) Two bananas is / #?are more than we need.

Notice that comparatives can be used in the prejacent of (3a) as seen in (20). This is a (weak) argument for taking the embedded clause in (3a) to in fact denote a degree.

(20) To get good cheese, you have to do more than go to the North End.

Then, let's assume than "(you) go to the North End", "you drink two beers", etc. each belong to a category of propositions for which a mapping to degrees is available. We denote this mapping as μ, and assume it is supplied by the context; the set of rankable propositions will be denoted asD_μ. Because universal degrees are available, μ does not have to correspond to a physical measurement. What we want is for the condition (3a) to mean "a degree as high as this one is reached". For the sake of simplicity (since syntax is not the focus of our discussion here) we can assume that this meaning obtains due to presence of the single operator AT-LEAST-DEG (given in (21)) in (3a), as in (22).

(21) $[\![\text{AT-LEAST-DEG } p]\!] = \lambda w.\, \exists q \in D_\mu,\ \mu(q) \geq \mu(p) \wedge q(w) = 1$

(22) To get good cheese, you only have to AT-LEAST-DEG [PRO go to the North End.]

The difference between AT-LEAST and AT-LEAST-DEG lies in the domain of propositions that they quantify over: instead of the formal alternatives to p, we now disjoin over the entire domain of μ, which is a potentially bigger set, and isn't constrained by the prosody. In (5), D_μ will be allowed to contain propositions like "you drink a glass of wine", and (16) will come out as equivalent to (17), thus letting us solve the problem of failure of necessity w.r.t. non-scalemates.[15]

We need some extra assumptions to get MS right. First, note that AT-LEAST-DEG immediately solves the prejacent problem: (3a) doesn't presuppose that you have to go to the NE, but that you have to do something that is ranked at least as high as going to the NE. In this respect AT-LEAST-DEG plays the same role as AT-LEAST in weakening the condition.

Do we get Minimality and Sufficiency? Going back for instance to the medals example: when (4) is asserted, assuming that it contains AT-LEAST-DEG and EXH, what we predict is that all outcomes that μ maps below a silver medal on the scale are ruled out. Furthermore, due to the presence of EXH, if any relevant outcome is mapped by μ strictly above a silver medal, then (4) implies that the outcome in question is above the minimal degree one needs to reach. What we want to predict is that getting a bronze medal is ruled out, but getting a gold or silver medal is not (Minimality) and that getting a gold medal is unnecessary (Sufficiency). It can be verified that the following condition is necessary and sufficient for our predictions and our desiderata to be the same:

(23) μ(you get bronze) $<$ μ(you get silver) $<$ μ(you get bronze)

This condition can be called a *monotonicity* constraint: the mapping to degrees has to be monotonic w.r.t. the scale, i.e. it has to respect the scale's ordering. It seems very natural

[15]Instead of making μ apply to propositions in a domain, we could have taken it to apply to worlds, or to situations. The general idea, however it is implemented exactly, is that the proposition being apparently expressed actually stands in for a gradable property. A very similar idea is proposed by Magri (2017, sec. 6) to account for the interaction between *only* and pragmatic scales.

that this condition should hold, for medals as well as for the other cases that we discuss.[16]

Thus, we have Minimality, Sufficiency, and Failure of Necessity w.r.t. scalemates. In this we only replicate what AT-LEAST already does. What is key is that AT-LEAST-DEG also allows for Failure of Necessity w.r.t. non-scalemates, because propositions not part of the scale might be in D_μ. Note that the monotonicity constraint, i.e. how μ ranks the scale, is μ's only specified property, and we do not assume that information about how μ ranks other elements of D_μ, and what D_μ even is, is available to participants in general.[17] In a sense, AT-LEAST-DEG weakens the condition so as to make only information about the scalar alternatives recoverable from an MS statement.

4.3.3 Are we really deriving Sufficiency?

There is a potential issue with the degree to which we have just weakened the condition: what we derive for Sufficiency might be somewhat weaker than what we want. Taking (3a) as an example, what we predict it to imply (putting together assertion and presupposition) is something akin to: you can do as much as go to the North End and get good cheese, and you can do less than go to South Boston and get good cheese. In other words, in some good-cheese worlds, the amount of effort you do is somewhere between going to the North End (included) and going to the South Boston (excluded). Notice that at no point are we specifically saying anything about going-to-the-North-End worlds here; i.e. this could very well be true even though there is no good cheese to be found in the North End, or anywhere else for that matter, and the only way to get good cheese is to order it online (which happens to be as much effort as going to the North End).

The issue is the following: AT-LEAST-DEG constitutes such a weakening that the propositional content that it embeds is lost. There is no role in our analysis for the set of worlds that the speaker chose to refer to to illustrate the degree they intended to communicate. To account for that, we would need to stipulate some sophisticated constraint, such as:

(24) μ, the scale and the goal are such that there is a threshold alternative p_0 on the scale such that the goal is attained if and only if a proposition whose degree is at least as high as the degree of p_0 is true.

Then, from Sufficiency we know that the threshold alternative is below South Boston: it is not necessary to reach the degree of South Boston. From Minimality we know that the threshold is at least as high as the North End: it is not sufficient to do any less than go to the North End. It follows that "you go to the North End" must be the threshold alternative, as it is the only alternative in the appropriate interval, and thus when you go to the North End, since "you go to the North End" is true and it ranks as high as the threshold alternative (i.e. itself), you can get good cheese. In addition to being unpleasantly convoluted, our constraint is not compositional, as it refers both to the goal and to the felicity of AT-LEAST-DEG involving a certain μ.

To avoid such a move, our only alternative is to accept that MS examples really have

[16]Note in particular that so far, we have treated pragmatic scales and contextual rankings as some kind of primitive object. Now that we have degree mappings at our disposal, we may keep our ontology lean by deriving pragmatic scales from them, and say that there is a pragmatic scale exactly when there is a salient mapping of a set of alternatives onto degrees. Then the monotonicity "constraint" is just a by-product of the definition.

[17]We could make this more explicit by making μ an existentially quantified variable over some class of natural metrics in the definition of AT-LEAST-DEG.

meanings as weak as we predict. This is perhaps not so implausible: there are other cases where utterances that we analyse to be truth-conditionally very weak lead listeners to strong inferences, based on world knowledge, typicality, relevance conditions etc.[18] It may be that the leap from our predicted weak sufficiency to the stronger form is an implicature of this kind, prompted by the following question: why would the speaker mention going to places if you could not get good cheese that way?

It must also be pointed out that the issue is in fact more general, and occurs with AT-LEAST already (thus it is not a problem specific to AT-LEAST-DEG). MS readings are easily accessible under "don't have to", even without *only*, as in (25).[19] Crucially, (25) implies that you can get good cheese in the North End. It would be natural to explain that as a scalar implicature: saying you don't have to do something implies that you can do it. However, one we've weakened "go to the NE" using even just AT-LEAST (which we need to do to account for the fact that (25) entails you can get good cheese somewhere closer), the scalar implicature we derive is "you can get good cheese by doing something at least as effortful as going to the North End", which doesn't entail that you can get good cheese in the North End.

(25) To get good cheese, you don't have to go to the North End.

We will leave open for now the question of how we can have our cake and eat it too, that is, weaken "go to the North End" in (25), while maintaining the inference that going to the North End (as opposed to doing anything of similar or higher effort) is in fact a way of getting good cheese. Hopefully whatever solution we find can also solve the problem of our overly weak Sufficiency condition.

4.3.4 Prosodic prominence, *only* and the structure of alternatives

Putting aside the problem described in the last section, we now have an analysis of MS whose key idea is to weaken the condition so much that no inference is recoverable about non-scale members. There are two things we still need to address: why the contrast between (3a) and the same sentence without *only*? And why is the marked prosody of (16) necessary?

Let's begin with the latter question. An interaction between the availability of MS readings and prosodic prominence is arguably expected under the analysis we have been developing. Indeed, following Rooth (1992), prosodic prominence is taken to constrain the alternatives set. If we identify the availability of AT-LEAST-DEG to that of a ranking of the

[18]A specific case of such inferences is the phenomenon known in the literature as *manner implicatures*, where truth-conditionally equivalent sentences may receive different interpretations based on specific choice of words.

[19]The MS reading of (25) that we are interested in is one where it licenses the inference that you can get good cheese closer than the NE, perhaps across the street. There is also a literal reading of (25) where it only means that you can get good cheese somewhere that is not the NE. Focus on *have* brings out this reading, as seen in (i).

(i) a. You don't HAVE to go the North End...
... You can also go to the corner store.
... You can also go to Switzerland.
b. You don't have to go to the NORTH END...
... You can also go to the corner store.
... ?? You can also go to Switzerland.

What remains also mysterious is why "go to the North End" can be interpreted as a scalar element in (ia), but not in (3b).

alternatives set, then prosody might prevent or force an MS reading in certain cases.

As we have already pointed out, such an interaction can be seen in (16), for which a certain, rather marked prosody is necessary for an MS reading to be accessible. This can be seen in (26).

(26) To be over the limit,

a. you have to drink TWO beers. You can also drink wine.

b. ?? you have to drink two beers. You can also drink wine.

c. #you HAVE to drink two beers. You can also drink wine.

Plausibly, the alternatives of (26a) are obtained by replacing the numeral with other numerals. We then obtain a scale that is totally ordered by entailment (if we assume, again, "at-least" semantics for numerals). However, in (26b), the alternatives are plausibly obtained by replacing "two beers" by other drinkable things, as in (27). The resulting set of alternatives is also ordered by entailment, but we do not have a total order this time. Instead, what we have a semi-lattice structure. Finally, in (26c), there is plausibly VERUM focus, or some other form of focus that results in an alternatives set similar to (28) for which no entailment-based order is available. Thus, it appears that the pattern in (26) is consistent with a constraint to the effect that whenever the alternatives are logically related, AT-LEAST-DEG can only occur if entailment defines a total order.[20]

(27) ALT("you drink [two beers]$_F$") = {"you drink x" | x is a plurality of drinks},

(28) ALT("VERUM$_F$ you drink two beers") = .
{"you drink two beers", "you don't drink two beers"}

Let us move to the second question. As we noted, (3b), without *only*, cannot receive a MS reading. Before we introduced the degree analysis, we took this to be a consequence of the fact that we do not easily accommodate a contextual ranking that would include "go to the North End", but that *only* forces us to. We can mostly maintain this view and rephrase it in terms of degrees: no degree mapping for "(you) go to the North End" is easily accommodated, but *only* has the effect of forcing us to accommodate that one exists anyway. We may adopt here the proposal made for independent reasons by Greenberg (2019) that *only* presupposes that a mapping to degrees is salient for its focus associate.

It should be noted that the effect of *only* interacts with the constraint on alternatives we just established: *only* may license an MS reading even for cases where one might think the alternatives form a semi-lattice that is not total, as in (29) (contrast with (26b)).

(29) You only have to drink two beers. You may also drink wine.

[20]As an aside, prosodic focus on bare numerals also leads to obligatory "exactly" readings for them in unembedded position, as seen in (i). It is tempting to relate this to the fact that (as we had to stipulate) exhaustification is obligatory in MS sentences: a generalization to the effect that exhaustification is obligatory in the presence of totally ordered alternatives would capture both facts.

(i) a. JOHN drank two beers. In fact, he drank four.

b. #John drank TWO beers. In fact, he drank four.

4.3.5 Disjunctive cases

To conclude this section, let us mention an interesting special case of MS, that this analysis extends nicely to: that of *disjunctive* MS, where a disjunction can be used to express MS along two dimensions at once. (30) is an example.

(30) To be over the legal limit, you have to drink TWO beers or THREE glasses of wine.

(30) appears to imply minimal sufficiency for both "two beers" and "three glasses of wine" (drinking that will put you over the limit, drinking less than that won't). It also demonstrates failure of Necessity, in the sense that it is compatible with the fact that drinking something else entirely (say, vodka) could put you over the limit. Failure of Necessity extends here to conjunctions of members of both scale: (30) can be true whether drinking one beer *and* two glasses of wine puts you over the limit or not; it implies nothing about what happens then.

Within our degree analysis, it seems natural to take (30) to have an LF equivalent at some level to (31). It can be straightforwardly verified that this delivers the right truth conditions.

(31) ..., you have to AT-LEAST-DEG [PRO drink two beers] or AT-LEAST-DEG [PRO drink three glasses of wine].

4.4 Getting MS by weakening modality

4.4.1 Better possibility and *have-to*

While the degree analysis of MS has been mostly satisfactory, we have been forced to rely on someone stipulatory constraints to account for the interaction with alternatives' structure, as described in Section 4.3.4. Besides, we saw in Section 4.3.3. that it wasn't entirely clear whether we hadn't derived a form of Sufficiency that is too weak. The latter issue stems from the radical and perhaps excessive weakening of the condition that the operator AT-LEAST-DEG performs. This might suggest a radically different approach: instead of weakening the condition through covert operators, we can make the modal itself directly sensitive to alternatives. This is in fact the only option that we have not explored yet: von Fintel and Iatridou (2007) have tried weakening *only*, and we have discussed in the previous section the option consisting in weakening the condition. Only the modal has been spared so far.

We explore here an approach consisting in giving weaker semantics to *have to*, specifically in the form of making it *alternative-sensitive*. In view of more general considerations on the alternative-sensitivity of modals (Villalta 2008), this line of analysis is quite natural. The role played by prosody in the cheese example, as discussed in Section 4.3.4, should also come out straightforwardly from an analysis of this sort. However, as we are going to see, the alternative-sensitive approach faces a number of challenges; we will not offer a solution to them here.

As already alluded to, the general idea of giving alternative-sensitive semantics to teleological *have-to* has already been proposed by Villalta (2008) and Krasikova (2010); the latter makes this move precisely to solve the "prejacent problem". Both proposals rely on the notion of "better possibility", which is due to Kratzer (1981).

Better possibility is defined as follows:

(32) Relative to a set of worlds Ω and an accessibility ordering $\leq$ on these worlds, a proposition p is a better possibility than a proposition q if:

a. For any world u in q, there is a world v in p such that $u \geq v$ (p is *at least as good a possibility* as q),

b. There is a world v in p such that for no world u in q do we have $v \geq u$ (q is not at least as good a possibility as p).

In particular, if the ordering is based on similarity to the actual world, this intuitively means that p is closer to the actual world than q, that is, that p-worlds are strictly more accessible than q-worlds. It can be checked that if q entails p, then p is at least as good a possibility as q (in particular, anything is at least as good a possibility as the contradiction, and the contradiction is at least as good a possibility as only itself).

Let us then give an alternative-sensitive semantics to *have-to*:

(33) *(to be refined)* "To p, have-to q" is true iff for any alternative q' to q (other than q itself), w.r.t. to a certain ordering, q is a better possibility than q'.

This is essentially what Villalta (2008) gives for *be necessary*, with the exception that we have left the ordering source unspecified. There are several natural choices. As Villalta (2008) does not discuss teleological modality, she proposes that the ranking should be based entirely on similarity to the actual world. Krasikova (2010) also takes the ordering to be similarity to the actual world, but takes the objects of comparison to be $p \wedge q$ and $p \wedge q'$ rather than q and q'. Her proposal is equivalent in many cases to having two ordering sources, a primary one based on p being true (where $g(w) = \{p\}$, using the notation of Kratzer (1981)), which we might call the teleological ordering source, and a secondary one based on similarity.[21] Finally, von Fintel and Iatridou (2005) propose to analyse teleological modality with strong modals (as opposed to weak modals such as *ought*) based only on the teleological ordering, with a modal base determined by physical possibility. We will adopt von Fintel and Iatridou's (2005) proposal here because it is simpler, and because the intuition that the actual world plays no particular role in determining the truth of teleological modal statements seems right; however our choice is not crucial as we will discuss.

If we adopt the teleological ordering, we can show that (33) is in fact completely equivalent to (34). Indeed, given our quasi-trivial order, if q' includes any p-world, then q' is at least as good a possibility as any other proposition. For q to be a better possibility than q', it has to be the case that q' should not be at least as good a possibility as q, and therefore that q' should be incompatible with p (within the bounds set by the modal base). Similarly, if q is incompatible with p, then anything other than a contradiction is at least as good a possibility as q; so again for q to be a better possibility than q', it has be the case that either q is compatible with p, or q' is a contradiction. Ignoring the pathological case where all the alternatives to q are contradictions, we can rephrase our statement as follows:

(34) *(simplified, to be refined)* "To p, have-to q" is true iff q is compatible with p, and for any alternative q' to q (other than q itself), q' is incompatible with p.

With a secondary ordering source based on similarity to the actual world or plausibility, the result we get is that either q is compatible with p and q' is not, or both q and q' are compatible with p, but $p \wedge q'$ is a more remote possibility than $p \wedge q$. We can note that, if q's only alternative is $\neg q$, then we are essentially asserting that p is incompatible with

[21] Cf. the discussion of systems of this kind by von Fintel and Iatridou (2005, sec. 6).

$\neg q$, or in other words that p entails q, over the domain of best worlds according to the secondary ordering source. This is more or less "normal" necessity.[22] We therefore still allow for "proper" necessity readings in cases where there is VERUM focus.

4.4.2 Adding effort to the ordering source: Krasikova's analysis of the pragmatic cases

Krasikova (2010) claims that the alternative-sensitive view of modality is sufficient to solve the prejacent problem of von Fintel and Iatridou (2007), as well as account for other MS examples not inolving *only*, involving both "pragmatic" scales (as in (4)) and the numeral scale (as in (5)). She argues, as we have done in Section 4.2.1, that scalarity is an essential feature of these examples. As we do, she assumes that, in (3a), the alternatives to "go to the North End" are a set of places you might think you could go to to buy cheese, and that a pragmatic ordering on these alternatives is contextually available.[23]

Second, she assumes that in the face of such pragmatically-ordered alternatives, the modality will be interpreted with respect to an ordering source that involves effort, or equivalently, it will be added to the goal that you want to minimize effort.[24] Thus, the prejacent of (3a) ends up paraphrasable as (35).

(35) To get good cheese and minimize your effort, you have to go to the NE.

(35) asserts that finding good cheese when ϕ_{NE} is true (you go to the NE) is a better possibility than for both ϕ_C (you cross the street) and ϕ_{SB} (you go to South Boston), the former because it is not plausible to find good cheese by crossing the street (this is essentially Minimality) and the latter because it is more effort to go to South Boston than to the North End. In effect, this analysis deals with failure of Necessity w.r.t. higher scalemates by assuming that we do in fact have Necessity, but not under the ordering source that we would expect.[25] Sufficiency also follows: if going to the North End is better than crossing the street even though the latter is more minimal, it has to be the case that going to the North End is compatible with getting good cheese.[26]

There are several concerns to find with this analysis. A first one is that the mechanism by which "minimalism" enters the ordering source is not clear. Presumably it is purely a matter of context. Notice, then, that we predict that depending on contextual variation of what is perceived as most effortful, MS examples could be used to indicate upper bounds on pragmatic scales rather than lower bounds. This might in fact be the case, as seen in (36).

(36) Context: *John is completely dominating his sport, and hasn't failed to win a race in years. He couldn't lose if he tried.*
John has to get a bronze medal in the semis to not make the finals.

[22] To be precise, if we drop the Limit Assumption, then defining necessity in our alternative-sensitive way and assuming VERUM focus ($\text{ALT}(q) = \{\neg q\}$) will yield a notion that Kratzer (1991) calls *weak necessity* and that is strictly weaker than the notion she calls necessity *tout court*.

[23] Krasikova (2010) considers that the prejacent problem, i.e. the lack of an MS reading for the cheese example without *only*, is purely due to the fact that one doesn't accommodate a context where an ordering is salient if that is not made obligatory by *only*; this is also how we have accounted for the contrast.

[24] If we analyse this interpretation as proceeding from an additional ordering source, then presumably this ordering source should rank after the teleological ordering source, but before considerations of plausibility in the hierarchy of orderings.

[25] Failure of Necessity w.r.t. non-scalemates is not dealt with, a point we will come back to.

[26] This is not quite logical Sufficiency, which we take to be desirable; cf. discussion in Footnote 29.

⇝ He'll make the finals iff he gets a silver or gold medal.

However, (36) and other attempts at getting upper-bound readings of teleological *have-to* have a certain emphatic flavour. Here, it seems to only be acceptable if we accept that John will never get anything below bronze (in which case there is no failure of Necessity), as is shown by the perceived incoherence of the discourse in (37a) (compare (37b)).

(37) a. Context: *same as above.*
#?John has to get a bronze medal in the semis to not make the finals. Perhaps he will not get a medal at all.
b. Context: *John is not particularly favourite in this competition.*
John has to get a silver medal in the semis to make the finals. Perhaps he will ?(in fact) get gold.

Thus, we might want our analysis not to allow these reversals, since the lower-bound examples do not have this emphatic flavour and do not set such strong conditions on the context. Besides, it seems problematic to assume that (3a) involves a ranking of worlds such that going to South Boston is considered remote, given the validity of the discourse in (38); if we assume that the possibility modal receives its standard semantics and that the necessity and possibility modals are evaluated relative to the same ordering, then (38) should be a contradiction.

(38) To get good cheese, you only have to go to the North End; you can also go to South Boston.

4.4.3 The case of the numeral scale

The problem becomes clearer when one at looks at numeral examples such as (39) (repeated from (5)). Krasikova (2010) assumes that numerals are generally interpreted under an "exactly" reading. Under the alternative-sensitive analysis, (39) then essentially asserts that drinking exactly two beers and being over the limit is a better possibility than being after the limit after having drunk exactly one beer or exactly three beers. This follows from the meaning of *have-to* adopted in (33) together with the natural assumption that the alternatives to the condition in (39) are those given in (40).

(39) To be over the legal limit, you have to drink TWO beers.

(40) $A = \{\text{"you drink } n \text{ beers"} \mid n \in \mathbb{N}\}$

Once again, we predict that reversal of the scale should be possible in appropriate contexts, that is, there should be a reading of examples like (39) asserting that to meet the goal, it is sufficient to drink two beers *or less than that*, in a context where drinking less is perceived as less of an effort. This prediction does not appear to be borne out, as seen in (41a) (compare (41b)). Even as part of a discourse that would make an "at-most" reading of the condition natural, the reading is not observed. Instead, inasmuch as the first sentence of (41a) is acceptable, it appears to presuppose that at least two beers will be drunk.[27]

[27]This is compatible with an analysis of the first sentence of (41a) where the numeral is interpreted as "exactly" and *have to* means regular Necessity. Indeed, in a context where it is presupposed that you will drink at least two beers, there is no failure of Necessity in this example.

(41) a. #To stay under the limit, you have to drink TWO beers. For instance, if you drink just one beer you will be under the limit.

b. To be over the limit, you have to drink TWO beers. For instance, if you drink three beers you will be over the limit.

The difficulty in "reversing the scale" makes another approach towards numeral examples attractive: if we instead assume that numerals are interpreted in an "at least" fashion, then we no longer expect the direction of the scale to be context-dependent. Indeed, as we have seen, even with a traditional, strong account of *have to*, we do not expect (39) to entail that drinking three beers does not put you over the limit as long as the numeral is interpreted as "at least two". In other words, the problem of failure of Necessity w.r.t. higher scalemates is at first glance avoided. We therefore do not need to postulate an additional mechanism that tinkers with the ordering source. Additionally, there are independent reasons to think that numerals are generally interpreted with an "at least" reading (cf. Spector 2013 for an overview of the issue).

The problem with the "at least" semantics for numerals is that it makes our analysis of *have to*, as given in (33) break down:

- We predict that drinking at least two beers is a better possibility than drinking at least one. Yet the altter proposition entails the former, and as already mentioned, if q entails q', then q' is always at least as good a possibility as q, and thus q is not a better possibility than q'. Then (39) should be a trivial falsehood.

- We predict that drinking at least two beers is a better possibility than drinking at least three. Given our world knowledge on how drinking relates to being over the limit, it is not plausible that drinking at least two beers should be compatible with being over the limit but drinking at least three should not. Thus, assuming a teleological ordering source, the only way this can be true is if the modal base excludes drinking at least three beers. With a secondary realistic ordering source, another way for it to be true would be if drinking at least three beers were a more remote possibility than drinking exactly two. There is no particular reason why either of these should be the case in many natural contexts where (39) is acceptable.

It is feasible to fix these two issues by tweaking the analysis. First, we want that the objects of comparison include some proposition equivalent to "you drink exactly one beer". This is indeed the set of worlds that we want to indicate represent a bad possibility, and not the set of worlds where you drink at least one beer. Second, we want that alternatives that lead to tautologies and contradictions not enter the computation. We may accomplish this by modifying our analysis of the teleological modal:

(42) *(refined)* "To p, have-to q" is true iff for any alternative q' to q, w.r.t. an ordering based on p being true, q is a better possibility than $q' \wedge \neg q$.[28]

[28] There are potential alternatives to this particular refinement that would do as well for our purposes. One consists in redefining the "at least as good a possibility as" relation so that it only quantifies over worlds that are in certain subsets of q and p. Kratzer (2012), for independent reasons, modifies the definition in this way, replacing q by $q - p$ and p by $p - q$ in the restrictor of the quantifiers. While her specific definition turns out to behave badly in our case, similar changes could allow us to stay closer to the "simple" version of the analysis outlined in (33). Alternatively, we could use exhaustification mechanisms to generate an object of comparison equivalent to "you drink exactly one beer" rather than build it in in the semantics. We discuss an idea along these lines in Section 4.4.4.

If we restrict our attention to the teleological ordering source, this can be simplified as follows:

(43) *(refined and simplified)* "To p, have-to q" is true iff q is compatible with p and for any alternative q' to q, $q' \wedge \neg q$ is not compatible with p.

With the new condition, the statement that (39) makes with respect to the n-alternative for $n \geq 2$ is trivial. Indeed, for $n \geq 2$, "you drink (at least) two beers" (q) is entailed by "you drink (at least) n beers" (q'), and therefore $\neg q \wedge q'$ is a contradiction, and as we already noted, any non-contradiction is trivially a better possibility than a contradiction. This is how we get failure of Necessity w.r.t. higher scalemates. As far as $n = 1$ is concerned, what (39) says is that "you drink two beers" is a better possibility than "you drink one beer but not two", i.e. that drinking at least two beers is compatible with being over the limit but drinking exactly one beer isn't. This is close to Minimality (though we will see that it is slightly wrong). As with the degree-based analysis, we can get Sufficiency through *exhaustification*: if the minimal amount of beers you had to drink was three, (39) would be true, but so would the same sentence with "three" instead of "two". As we desire, we do not derive Necessity: all what we know about drinking (exactly) two beers is that it is *compatible* with being over the limit, leaving open whether drinking something else entirely is a possibility.[29]

As an interim conclusion to our discussion of the numeral cases, what we have seen here is that the alternative-sensitive approach can deliver failure of Necessity w.r.t. non-scalemates and Sufficiency in two ways. If, following Krasikova (2010), we assume "exactly" semantics for the numeral, and we allow for some notion of minimalism to make it to the ordering source, we make the unwelcome prediction that the scale could be reversed in the appropriate context. If we instead adopt "at least" semantics for the numeral, and we modify the specific analysis of teleological modals, we can avoid the problematic prediction, although the entry for the modal that we end up with is perhaps *ad hoc*. Under either approach, the issue described in Section 4.3.3 does not occur: we have kept the condition strong enough that we are still clearly talking about beer-drinking worlds, and our (enriched) semantics properly entails that drinking exactly two beers is compatible with being over the limit.

4.4.4 Why we are still getting MS wrong

Unfortunately, the analysis of the beer example (5)/(16)/(39) we have been developing in this section is faced with the following major problem, in both of its incarnations: it does not actually deliver failure of Necessity w.r.t. non-scalemates. Indeed, we end up predicting

[29] One may be find it concerning that we do not derive logical sufficiency either: our semantics does not entail that drinking exactly two beers *will* reliably put you over the limit. We argue that this result is correct. In this particular example, it proceeds from world knowledge that, if drinking two beers can put you over the limit, then it will generally do so. In the more general case, MS does not have to come with logical sufficiency. (i), for instance, exhibits failure of necessity in the sense that it may accurately describe a situation where students who do a presentation do not have to submit anything to pass. It also exhibits failure of (logical) sufficiency in the sense that it does not necessarily imply that students who submit two squibs will pass. It may be that students who pass through submitting squibs, rather than doing a presentation, also have attendance requirements. It may also be that the squibs have to be up to a certain standard. What (i) does imply is that students do not need to submit three squibs, and that submitting one squib won't help them pass, which we take to be all what MS means.See also Section 4.1 of von Fintel and Iatridou 2007 for a closely related point.

(i) Students have to submit TWO squibs to pass.

that drinking exactly one beer is incompatible with being over the limit, where "exactly one beer" means "one beer but not two beers". However, this cannot be true in any sensible context: someone may very well be over the limit after having drunk one but not two beers, as long as they have been drinking wine as well, and we may still judge (39) true.

Put another way, the problem is that the worlds we wanted to rule out were those where you drink exactly one beer *and no wine*. Instead, we ruled out worlds where you drink one beer and not two.

A somewhat natural move consists in assuming that the domain of worlds that the modal quantifies over is reduced to worlds where nothing else than beer is drunk. This could be taken to be a case of loose association with focus as described by D. I. Beaver and Clark (2008). The embedded clause's focus structure is anaphoric on a salient situation where beer is drunk (as this is what the disjunction of all alternatives amounts to). The modal's domain variable (or its implicit modal base, as one will have it) is identified to the corresponding set of worlds containing a situation of this kind through a higher-level pragmatic process.

It should be noted that the focus structure we're dealing with does not in fact pre- suppose or suggest in any way that no wine is drunk. (44) is perfectly compatible with John having drunk some wine, and does not even suggest otherwise. On the contrary, narrow focus suggests that John may very well have drunk wine, but that whether he did is irrelevant for the speaker's purposes.

(44) John drank TWO beers yesterday.

Thus, if we rely on domain restriction, we cannot say that it proceeds from loose association with focus. We have to assume that non-systematic domain restriction occurs until the utterance is not trivially false, with worlds where no wine is drunk being perhaps more salient, accessible or prototypical than other beer-drinking worlds.[30]

4.4.5 Local conclusion

Let us sum up our discussion of the alternative-sensitive approach. We can implement it in two ways: either we let some consideration of minimal effort, or something of this kind, enter the ordering source, and we adopt "exactly" semantics for numerals, as Krasikova (2010)

[30]It may appear that once we have such domain restriction, failure of necessity is no longer a problem. Indeed, because we are only quantifying over worlds where no wine is drunk, it becomes actually true that you have to drink two beers to be over the limit. Then, since the fact that this was false was our motivation for introducing AT-LEAST-DEG or changing the semantics we assign to the modal, it appears that there is nothing remarkable going on any more in the beer case, and that for the other examples we may stick to the analysis of Alonso-Ovalle and Hirsch (2018). Here is the problem: we already had to stretch domain restriction beyond what prosodic presuppositions have to offer. Extending a domain-restriction analysis to cases of disjunctive MS is going to require an even more *ad hoc* notion of domain restriction.

Recall that disjunctive MS consists in cases where a disjunction can be used to express MS along two scales at once, as in (i) (repeated from (30)).

(i) To be over the legal limit, you have to drink TWO beers or THREE glasses of wine.

Crucially, (i) does not specify whether drinking one beer *and* two glasses of wine puts you over the limit or not. Then, we want to analyse it in such a way that we do not predict that a world where you have drunk both beer and wine is ruled out. If all we have is domain restriction, we are going to need the domain to be a set of worlds where beer or wine are drunk, but not both at the same time. It is implausible that this disconnected domain would be implicitly accommodated through a pragmatic process.

does, or we slightly complicate the modal and adopt "at least" semantics for numerals.[31] In both cases, we account for failure of Necessity w.r.t. higher scalemates and for the appropriate notion of Sufficiency (the condition stated is compatible with the goal). In both cases, we avoid the problem described in Section 4.3.3. In only the former case, we make a possibly unwanted prediction that the direction of the scale can be reversed by context. Most importantly, in both cases, we derive a notion of Minimality that is too strong: we predict that the "bad" propositions and the goal are actually incompatible, even though we have seen that this is not actually tenable in our examples. We are essentially missing failure of Necessity w.r.t. non-scalemates.

We might try to weaken the predicted meaning by assuming that considerations of physical likelihood or typicality can affect the ordering source or induce domain restriction. However, in our examples, the situations we want to restrict ourselves to (drinking beer but nothing else, submitting squibs but not doing a presentation) do not seem to be especially likely or prototypical, and as far as I can tell, the context dependency that one would expect if likelihood entered into consideration is not there.[32]

Thus, our problem does not appear to be easily solved through established views on how modals might be alternative-sensitive, which tends to suggest that the "second route", where we weaken the embedded clause, was the more promising one. However, it might be the case that another implementation of the alternative-sensitive intuition, possibly based on a different definition of better possibility, would prove more successful and let us avoid the problems that we identified for the degree-based account.

4.5 Issues with the presupposition of *only* and the status of Minimality

The issue we discuss in this section is a puzzling feature of MS examples using *only*; since the rest of our discussion mostly focussed on examples without *only*, this section is relatively orthogonal to it. It does, however, make the various theories come apart on a specific point: whether Minimality is linked to the prejacent inference of *only*.

Recall that we stated that the cheese example (3a) implicates in some way that you cannot get good cheese across the street, a meaning component we call Minimality, and that we took this fact as evidence against the decompositional analysis due to von Fintel and Iatridou (2007). In both the analysis based on degrees, and, *a priori*, an analysis along the alternative-sensitive line, Minimality is an entailment of MS statements without *only*. Thus, Minimality in MS statements with *only* can be seen as an instance of the inference to the prejacent. If, following Horn (1969), we take the inference to the prejacent to be a presupposition, then Minimality should be a presupposition. In contrast, under the decompositional

[31] The second option can be extended to pragmatic cases if we assume that AT-LEAST is present in the relevant examples.

[32] Possibly, a solution would involve the notion of minimal situations in the sense of situation semantics (Kratzer 2019). Intuitively, the reason that, for instance, in the scenario described in Footnote 29, (i) is true is that a greater share of the sufficient conditions for passing are met when you write two squibs than when you write just one. In particular, if we know that a student wrote two squibs, the information that they meet the attendance requirements lets us conclude that they will pass, while otherwise it does not. From this intuition we can derive a notion of comparative possibility over situations, where a situation is better than another in view of a goal if adding extra information to it will let us conclude that the goal is attained in more cases. Unfortunately, as far as I can tell, this approach does not lead to a notion of alternative-sensitive necessity that reduces to the traditional view in simple cases, which seems like a problem.

analysis as originally proposed by von Fintel and Iatridou (2007), an existential presupposition conspires with the at-issue meaning of *only* to produce the inference to the prejacent in the general case. In an MS statement, this does not occur and Minimality is not generated at all. If we adopt the "fixed" decompositional analysis as described in Section 4.2.2, Minimality will be a presupposition that is triggered below the modal. An alternative view truer to the spirit of von Fintel and Iatridou's would consist in generating Minimality as an implicature: to do so, we can take (3a) only entails Sufficiency and has a trivial presupposition, for instance through the entry for QUE in (45), and additionally assume that it has as an alternative (46). When (3a) is uttered, a scalar implicature to the effect that (46) is false is triggered, which is how Minimality obtains.

(45) $[\![\text{QUE}\ p]\!] = \lambda w : [\exists p' \in \text{ALT}(p),\ p'(w) = 1].\,[\exists p' \in \text{ALT}(p),\ p' \succ p \wedge p'(w) = 1]$[33]

(46) To get good cheese, you only have to cross the street.

To summarize, under both the degree-based and the alternative sensitive approaches, Minimality in (3a) proceeds from the inference to the prejacent (traditionally seen as a presupposition). In the decompositional analysis, Minimality will not exactly be an instance of the inference to the prejacent, and in particular we can generate it as a scalar implicature under a specific version of the analysis.

In view of the debate on *only*'s prejacent, it is therefore interesting to look at properties of the Minimality inference in MS examples, and see whether we can ascertain its status as a presupposition or an implicature. The main fact speaking in favour of the view that Minimality is a presupposition is that there is a contrast between the Minimality inference that MS statements trigger, and the similar inference deriving from an actual statement of sufficiency using *enough*, as in (47a). To begin with, we already discussed the fact that (3a) is somewhat degraded in a context where you can in fact get good cheese across the street. It appears there is a contrast with (47a), which seems to be more universally accepted (from an informal survey). In a similar fashion, the "enough" statement might be more easily defeasible, as seen in (48), though the judgement is not clear.

(47) Context: *you can in fact get good cheese in the neighbourhood.*
  a. ? To get good cheese, it is enough to go to the North End.
  b. #? To get good cheese, you only have to go to the North End.

(48) a. ? To get good cheese, it is enough to go to the North End. In fact, you can also find some across the street.
  b. ?? To get good cheese, you only have to go to the North End. In fact, you can also find some across the street.

The two constructions also contrast in questions, where MS Minimality appears to project, as judged by the degraded status of (49b), while "enough" Minimality disappears. All this points to an analysis of the Minimality inference triggered by *enough* statements as a *bona fide* scalar implicature: it is deafeasible and disappears in questions. In contrast, the MS statement is of dubious defeasibility, and projects from questions, suggesting it is a presupposition. All this would speak against the view that Minimality is an implicature, and therefore against the decompositional account as outlined above.

(49) I have no idea where you can get good cheese in Boston.

[33]Note the use of $\succ$ rather than $\succeq$.

a. Is it enough to go to the North End?
b. #Do you only have to go to the North End?

There are, however, also arguments in favour of the implicature analysis. To begin with, as we have already mentioned, the Minimality inference still seems to be easily defeasible, as speakers surveyed generally accept (48b) even when they perceive a contrast. Presuppositions are generally not defeasible, at least when they are generated in unembedded environments (presumably because what is presupposed is also entailed). Note also that in the degree-based analysis and the alternative-sensitive analysis, Minimality proceeds from the prejacent inference; yet as discussed by D. I. Beaver and Clark (2008, chap. 9), the prejacent inference of *only* is not cancellable in simple positive sentences like (50).[34]

(50) Muriel only likes Hubert. (#In fact, she doesn't like Hubert.)

Another puzzling fact is that when an MS statement with *only* occurs in the antecedent of conditionals, Minimality does not constrain either the overall context or the antecedent. The reported and predicted behaviour of presuppositions (cf. e.g. Heim 1983) is that they should "project" from the antecedent of a conditional, that is to say, they should constrain the overall context. One may easily construct examples such as (51) where the prejacent presupposition of *only* fails to project, and is instead accommodated into the antecedent (as if it were entailed): (51) asserts that Muriel will accept if the world is such that she likes Hubert (prejacent) and noone else (assertion). In similar MS cases, Minimality neither projects (as we might expect, if it is a presupposition) nor is locally accommodated (as we might expect, if it is a prejacent inference); it just disappears. Example (52a) does not imply that you in fact have to drink two beers to go to jail (this would make it a contextual contradiction), and it also does not imply that (52b) is true. (52b) is in fact judged to be a plain falsehood. This is exactly what one would expect under the implicature analysis: implicatures are known to disappear in the antecedent of a conditional (because it is a downwards-entailing environment). In fact, we observe exactly the same thing when we look at the "enough" statement and its well-behaved Minimality implicature.

(51) If Muriel only likes Hubert, she will accept, but I doubt she likes Hubert.

(52) a. If you only have to drink two beers to go to jail, then you might still go to jail after just one beer.
b. #If you have to drink two beers to go to jail, then you might still go to jail after just one beer.
c. If it is enough to drink two beers to go to jail, then you might still go to jail after just one beer.

(Brian Buccola, p.c.)

To summarize, the Minimality inference in MS contrasts with the well-behaved implicature of *enough* statements. This first set of facts suggest it should be seen as a presupposition, which could be derived from a presuppositional prejacent inference. Yet, it also exhibits certain behaviours that are abnormal for a presupposition, and in this respect it does not behave like the prejacent inference in simpler sentences. This second set of facts speaks in favour of a view of Minimality as an implicature and is compatible with a version

[34] D. I. Beaver and Clark (2008) in fact argue that the defeasibility patterns of the prejacent inference in unmodalized positive and negative sentences are fully consistent with neither an analysis as a presupposition nor as an implicature, but that they can be understood under a refined version of the presuppositional view.

of the decompositional analysis. MS statements with *only* therefore constitute a challenge for analyses of MS along all routes discussed in this article, as well as for various views on the prejacent inference of *only*. We do not have a solution to this puzzle to offer here.

4.6 Conclusion

This article discusses three routes towards accounting for MS statements with minimal disruption to Montagovian composition. They consist in, respectively, weakening *only* (the decompositional analysis), weakening the condition (the degree-based analysis), and weakening the modal (the alternative-sensitive approach). All three routes have arguments in their favour *a priori*. Yet, they quickly run into a number of challenges.

The decompositional analysis ties in well with other observations of unexpected weakness of *only* (Horn 1996, a.o.) as well as interactions between modals and negative elements (von Fintel and Iatridou 2007). Yet, the weakening of *only* might be problematic in some other ways (Alonso-Ovalle and Hirsch 2018). Additionally, in the implementation of von Fintel and Iatridou (2007), the Minimality inference is wholly absent and at best seen as a side-effect of another component, Easiness, whose origin is not accounted for. Alternative implementations can reintroduce Minimality proper, but only if we are willing to make the novel assumption that silent comparative elements exist. Finally and most problematically, MS statements not involving *only* (where Minimality is clearly part of the truth conditions) are unexplained.

The degree-based approach "keeps *only* strong" (Alonso-Ovalle and Hirsch 2018), does not require any unusual assumption about the composition of the sentence, and can account for MS statements without *only*. Its main defect is that it results in extremely weak truth conditions for MS statements, from which we cannot recover our intuition as to what Sufficiency is (Section 4.3.3). This gap can plausibly be explained away by invoking manner implicatures or "leaving it to the pragmatics" in some other way.

The alternative-sensitive approach is very promising *a priori* in that it builds up on a more general account of alternative-sensitivity in modals (Villalta 2008). It also keeps *only* strong and eschews the main problem of the degree-based approach. However, if paired with "exactly" numerals, it makes the incorrect prediction that the direction of the scale is a matter of context. If paired with "at least" numerals, it does not weaken the truth conditions enough, and we still predict that MS statements should be plain false in contexts where we judge them to be true.

Finally, the fact that in MS examples involving *only*, the Minimality inference does not behave exactly like a presupposition nor like a scalar implicature, is problematic for all approaches, inasmuch as they can only make one or the other prediction.

In the end, the issue of how to integrate MS statements into our understanding of modal semantics remains as pressing as ever.

References

Alonso-Ovalle, Luis and Aron Hirsch (2018). "Keep *only* strong". In: *Semantics and Linguistic Theory 28*. Ed. by Sireemas Maspong et al., pp. 251–270.

Bale, Alan C. (2008). "A universal scale of comparison". *Linguistics and Philosophy* 31.1, pp. 1–55.

Beaver, David I. and Brady Z. Clark (2008). *Sense and sensitivity: How focus determines meaning*. Oxford: Wiley-Blackwell.

Coppock, Elizabeth and David Beaver (2014). "Principles of the Exclusive Muddle". *Journal of Semantics* 31.3, pp. 371–432.

von Fintel, Kai and Sabine Iatridou (2005). "What to do if you want to go to Harlem: anankastic conditionals and related matters". Ms. Massachussetts Institute of Technology.

von Fintel, Kai and Sabine Iatridou (2007). "Anatomy of a modal construction". *Linguistic Inquiry* 38.3, pp. 445–483.

Greenberg, Yael (2019). "A scalar presupposition with universal force: *only* as (only) a scalar mirror image of *even*". Ms., Bar Ilan University.

Groenendijk, Jeroen and Martin Stokhof (1984). "Studies on the semantics of questions and the pragmatics of answers". PhD thesis. University of Amsterdam.

Heim, Irene (1983). "On the projection problem for presuppositions". In: *Formal Semantics: The Essential Readings*. Ed. by Paul Portner and Barbara Partee. Blackwell, Oxford, pp. 249–260.

Horn, Laurence R. (1969). "A presuppositional analysis of *only* and *even*". In: *Papers from the fifth regional meeting of the Chicago Linguistic Society*. Ed. by Robert I. Binnick et al., pp. 98–107.

Horn, Laurence R. (1996). "Exclusive company: *Only* and the dynamics of vertical inference". *Journal of semantics* 13.1, pp. 1–40.

Klinedinst, Nathan (2004). "Only scalar *only*". Handout.

Krasikova, Sveta (2010). "Sufficiency inference in anankastic conditionals". In: *Proceedings of SALT 20*. Ed. by Nan Li and David Lutz, pp. 91–108.

Kratzer, Angelika (1981). "The notional category of modality". In: *Words, worlds, and contexts: New approaches in word semantics*. Ed. by Hans J. Eikmeyer and Hannes Rieser. De Gruyter, pp. 38–74.

Kratzer, Angelika (1991). "Modality". In: *Semantics: An international handbook of contemporary research*. Ed. by Arnim von Stechow and Dieter Wunderlich. Berlin: de Gruyter, pp. 639–650.

Kratzer, Angelika (2012). "The Notional Category of Modality". In: *Modals and Conditionals: New and Revised Perspectives*. Oxford University Press.

Kratzer, Angelika (2019). "Situations in Natural Language Semantics". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2019.

Magri, Giorgio (2017). "Blindness, Short-sightedness, and Hirschberg's contextually ordered alternatives: a reply to Schlenker (2012)". In: *Linguistic and psycholinguistic approaches on implicatures and presuppositions*. Ed. by Salvatore Pistoia-Reda and Filippo Domaneschi. London: Palgrave Macmillan, pp. 9–54.

Rett, Jessica (2014). "The polysemy of measurement". *Lingua* 143, pp. 242–266.

Rooth, Mats (1992). "A theory of focus interpretation". *Natural Language Semantics* 1.1, pp. 75–116.

Schwarz, Bernhard (2005). "Scalar additive particles in negative contexts". *Natural Language Semantics* 13.2, pp. 125–168.

Spector, Benjamin (2013). "Bare numerals and scalar implicatures". *Language and Linguistics Compass* 7.5, pp. 273–294.

Villalta, Elisabeth (2008). "Mood and gradability: an investigation of the subjunctive mood in Spanish". *Linguistics and Philosophy* 31.467.

Chapter 5

A decision-theoretic perspective on the square of Aristotle

Written with Benjamin Spector (IJN, ENS/EHESS/CNRS). This chapter was originally published as:

> Émile Enguehard and Benjamin Spector (2021). “Explaining gaps in the logical lexicon of natural languages. A decision-theoretic perspective on the square of Aristotle”. *Semantics and Pragmatics*.

Abstract

Across languages, certain logically natural concepts are not lexicalized, even though they can be expressed by complex expressions. This is for instance the case for the quantifier *not all*. In this paper, we propose an explanation for this fact based on the following idea: the logical lexicon of languages is partly shaped by a tradeoff between *informativity* and *cost*, and the inventory of logical expressions tends to maximize average informativity and minimize average cost. The account we propose is based on a decision-theoretic model of how speakers choose their messages in various situations (a version of the *Rational Speech Act* model).

Acknowledgements

We thank Emmanuel Chemla and Isabelle Dautriche for useful discussions, as well as three anonymous reviewers for *Semantics & Pragmatics* for their helpful suggestions and remarks.

5.1 Introduction

The Aristotelian square of opposition consists of the following four categories of logical statements:

1. *Universal (A):* All As are Bs.

2. *Negative universal (E):* No As are Bs.

3. *Existential (I):* Some As are Bs.

4. *Negative existential (O):* Some As are not Bs / Not all As are Bs.

An observation due to Horn (1973) is that English can express A, E and I more concisely than it can express O. Specifically, English can use a quantifying determiner to express A, E and I statements, but there is no corresponding word for O statements; instead one must use an additional negation. This can be seen in (1).

(1) a. *All* As are s. / *Every* A is a B.
b. *No* As are Bs.
c. *Some* As are Bs.
d. **Nall* As are Bs. / **Nevery* A is a B.

The general conclusion is that the abstract logical operators corresponding to A, E and I are (usually) *lexicalized* while O never is. This observation can be generalized to other languages as well as to temporal quantifiers and modal operators. For instance, 'not always' is less likely to be lexicalized than 'sometimes', 'never' and 'always' across languages, and 'unnecessary', which is lexicalized in English, is across languages less often lexicalized than 'possible', 'necessary' and 'impossible'.

At first sight, this observation is somewhat surprising, since the four operators all satisfy familiar constraints on possible quantifiers: not only conservativity, but also, in this case, monotonicity.[1] Furthermore, while O and A are each other's negation, and one might think that this would make having both unnecessary, I and E are each other's negation. Similarly, while I is the dual of A, O is the dual of E. The difference in terms of lexicalization between I and O is therefore unexpected.

Horn (1973) suggested an explanation why only three operators are needed: because they are logically related to one another (A entails I, E entails O, A and O and I and E are one another's negation), the four operators only allow us to distinguish between three basic situations or categories of worlds:

(2) $w_{\forall}$: Worlds where all As are Bs (A and I are true, E and O are false);
$w_{\exists\neg\forall}$: Worlds where some but not all As are Bs (I/O are true, E/A false);
$w_{\neg\exists}$: Worlds where no As are Bs (E and O are true, A and I are false).

Furthermore, the mechanism of *scalar implicatures* lets the speaker indicate in which of these situations we are with just three operators. A and E statements can be used to indicate that we are in $w_{\forall}$ and $w_{\neg\exists}$ respectively. Then, if the speaker utters I, the listener may reason that A is false, as otherwise the speaker could have been more informative by saying A. For

[1] Barwise and Cooper (1981) suggested that conservativity and monotonicity are semantic universals. See also Chemla, Buccola, and Dautriche (2019) for a recent discussion of a number of semantic properties, including monotonicity, which make some concepts more likely than others to be lexicalized across languages.

instance, an *I* statement like (3) is understood in most contexts to imply that the stronger *A* statement in (4) is false, or, equivalently, that we are in $w_{\exists\neg\forall}$; this is a well-known instance of the scalar implicature of an *I* statement. Thus, with just three operators, we are able to refer to the 3-way partition of worlds outlined in (2) with maximal precision; adding *O* to the mix would not allow us to be any more informative.

(3) Some of my colleagues are nice people.

(4) All of my colleagues are nice people.

Horn's idea, however, does not explain why, in a given family of logical operators, it is always the set $\{A, E, I\}$ which is lexicalized, and not the other 3-element set $\{A, E, O\}$, with which a symmetric argument could be made. Indeed, *O* statements also trigger an inference to the effect that the corresponding *E* statement is false. Furthemore, as we will discuss, speakers sometimes do use *O* statements, and there are contexts where *O* statements are more felicitous than *I* statements (cf., e.g., the contrast in (12) below).

Horn (1973) proposes that *E* and *O*, being negative (downward-entailing), are marked in some sense and therefore dispreferred. Katzir and Singh (2013) generalize Horn's idea by proposing that at a certain level, logical operations are expressed in terms of certain primitives; this has the consequence that *E* and *O* have more complex representations. This is in line with findings suggesting that monotone-decreasing operators are harder to process than monotone-increasing ones (Geurts and van der Slik 2005).

Both Horn (1973) and Katzir and Singh (2013) essentially attempt to break the symmetry between *I* and *O* by assuming that *O* is inherently marked in some sense, making a lexicon that includes *O* rather than *I* dispreferred. Katzir and Singh (2013) implement this idea by assuming a specific inventory of cognitive primitives such that monotone increasing operators have simpler representations. We should note however, that in the absence of such hypotheses about cognitive primitives, there is no reason to believe that *I* is intrinsically computationally less complex that *O*. Consider for instance a measure of complexity based on the semantic automata that can represent a given quantifier (see, e.g., Steinert-Threlkeld and Szymanik 2019; Katzir, Lan, and Peled 2020). An automaton for *I* takes a sequence of 0s and 1s and returns 'true' if the sequence contains at least one occurrence of 1. It is trivial to convert such an automaton into one with exactly the same structure which returns 'true' if the sequence contains at least one occurrence of 0, thus encoding the meaning of *O*. As to the observed cognitive cost of sentences containing negation and other monotone-decreasing operators, it should be noted that such sentences tend to be *syntactically* more complex than their 'positive' counterparts, so that we should not conclude that this cost is specifically tied to the logical meaning of the relevant operators. We discuss these points with more details in Section 5.4.2, in connection with recent proposals that are to some extent related to ours.[2]

We approach the problem from a very different angle. First we will argue that there are principled reasons why, irrespective of which corners of the Aristotelian square are lexicalized, *O*-statements are expected to be less frequently used than *I*-statements. Given the well-established relation between frequency and lexicalization (the more frequently an expression is used, the more likely it is to be lexicalized), this might provide an explanation for why *I* tends to be lexicalized but *O* does not. Second, we will show that under certain

[2]See, among others, Penka (2011), Zeijlstra (2011), and Buccola and Spector (2016), for arguments that monotone-decreasing quantifiers such as *no* or *fewer than 10* are to be syntactically decomposed into a negation-like operator and an upward-entailing quantifier, based in part on the availability of so-called 'split-scope' readings.

plausible assumptions, lexicalizing $\{A, E, I\}$ is optimal compared to $\{A, E, O\}$, in that it maximizes the *expected utility* that speakers can receive from using the language, where the utility of a single message in a single occasion of use depends on a trade-off between how informative the message is in this situation and how costly it is. The account we will propose is in line with the view that a number of features of natural languages can be understood as maximizing the overall utility of a language (cf. Gibson et al. 2019 and the references cited therein). We will make use of the same information-theoretic definition of utility that is used in the Rational Speech Act model of pragmatics (RSA; Goodman and Stuhlmüller 2013, and Bergen, Levy, and Goodman 2016, where a cost-term is introduced in the utility function of messages), but our account is otherwise not couched in a game-theoretic framework. For both lines of explanation, a crucial ingredient of our account is the observation (already made in Chater and Oaksford 1999) that, on average, an O-statement (e.g., 'Not all of the guests were drunk') is less informative than its corresponding I-statement (e.g. 'Some of the guests were drunk').

On our approach the asymmetry will be derived directly from the truth conditions of the operators, together with independently motivated assumptions about the meanings of lexical predicates and general principles of language use. In particular, cognitive or morphological primitives will play no role in our explanation.[3]

We will proceed as follows: in Section 5.2, we will explain why I-statements are expected to be used more frequently that O statements. In Section 5.3, we provide a model of the expected utility of a lexicon, in which the expected utility of a lexicon based on $\{A, E, I\}$ has a greater expected utility than one based on $\{A, E, O\}$.

5.2 Why *I* is expected to be more frequent than *O*

It would be no surprise to find, in a corpus study, that I-sentences of the form *Some of the NPs VP* (e.g., *Some of the guests were drunk*) have on average more occurrences than corresponding O-sentences of the form *Not All of the NPs VP* or *The NPs AUX not all VP* (e.g., *Not all of the guests were drunk*, *The guests were not all drunk*). Given that, in English and in other languages, O is not lexicalized, O-sentences are bound to be syntactically more complex than I-sentences, which would be enough to explain the observed difference in frequency of occurrences, on the plausible assumption that, everything else being equal, more complex constructions are less frequent than less complex ones. Such a corpus-based observation would not provide any convincing explanation for the fact that I is lexicalized and O is not, since the lexicalization facts might in fact explain the observed frequencies.

What we want to argue here is that even if the four corners of the Aristotelian square were lexicalized, O-sentences would still be less frequently used than I-sentences.

Imagine a speaker who knows that, say, some but not all of the guests at a certain party were drunk. The two 'universal' corners of the square cannot be used, as they yield false statements. If the speaker is to use one of the four quantifiers of the square, the choice to make is, then, between the following sentences:

(5) a. Some of the guests were drunk.

[3] Several recent works (Denić, Steinert-Threlkeld, and Szymanik 2021; Uegaki 2020, a.o.) also explore the inventory of lexical quantifiers from an information-theoretic perspective, but their approach relies on specific assumptions on lexical primitives or meaning spaces similar to those made by Katzir and Singh (2013). We will discuss the relevance of these analyses to Horn's puzzle and their similarities and differences with our proposal in Section 5.4.2.

b. Not all of the guests were drunk.

Now, both sentences trigger a 'some but not all' implicature, so once this is taken into account they should be equally effective. If we only consider the literal meaning of both sentences, and use a notion of informativity based on entailment, neither sentence is more informative than the other, since there is no entailment relation in either direction. On such a view of informativity, Grice's maxim of Quantity cannot help us choose between the two sentences. However, suppose that we use instead a probabilistic notion of informativity, where, in a given context, a proposition ϕ is more informative than a proposition ψ if, before she hears the sentence, the hearer considers ϕ less likely than ψ. In that case, we have to ask which of the two propositions in (5) is the least likely to be true in the context of utterance, from the point of view of the listener. In many contexts, the probability that no guest is drunk (i.e., the probability of the negation of (5a)) is higher than the probability that all guests are drunk (i.e., the probability of the negation of (5b)). Equivalently, in such contexts, (5a) is less likely to be true than (5b), hence, given the perspective we adopt here, more informative. Assuming that speakers tend to prefer more informative sentences, (5a) will be used in a such a context.

Suppose now that, in most contexts, I-statements are more informative than O-statements in this sense. Then we would predict that, irrespective of lexicalization, I would be more frequently used than O. Such a prediction might in turn help explain the lexicalization facts.

To substantiate an account of this type, we need to a) establish that indeed the choice between I and O is partly governed by probabilistic informativity, and b) that there are good reasons to think that across contexts I is indeed more informative (in this sense) than O.

Before turning to these issues, we will provide a model of message choice which captures the reasoning we have just sketched, inspired by the RSA model of pragmatics (Goodman and Stuhlmüller 2013).

5.2.1 A simple model of the pragmatics of the Aristotelian square

We assume that there is a certain set of possible worlds Ω, and a set of messages $\mathcal{M}$. The speaker knows exactly what the world is, while the listener's prior beliefs are represented by a probability distribution over Ω. The listener's prior beliefs are known to the speaker and more generally part of the Common Ground. Upon hearing a message, the listener updates their belief distribution. Thus, a model of the listener's behaviour is a function $L(w|m)$ giving the probability the listener assigns to world w after having heard message m. One particular listener behavior is that of a *literal listener*. The literal listener L_0 has a prior distribution P_0 over worlds. They also have a notion of the semantics of each message: to each message m, they assign a set of worlds $[\![m]\!]$ where the message is true. Upon hearing m, the listener conditionalizes their belief distribution on m being true:

(6) $$L_0(w|m) = P_0(w|[\![m]\!]) = \begin{cases} 0 & \text{if } w \notin [\![m]\!], \\ \frac{P_0(w)}{P_0([\![m]\!])} & \text{if } w \in [\![m]\!]. \end{cases}$$

We assume for the time being that the speaker chooses her message m in the following manner: if they are in world w, they pick the message that maximizes the probability that the listener will assign to w after processing it.[4] We write $S(w; L)$ to refer to the message

[4]Here we depart from the standard RSA model, where the speaker is not fully rational, and does not

chosen by a speaker who believes w.[5]

(7) $S(w;L) = \arg\max_m P_0(w|m)$

Now, let us assume that worlds are individuated by whether all ($w_\forall$), some but not all ($w_{\exists\neg\forall}$), or no ($w_{\neg\exists}$) guest is drunk. Let us assume that there are four messages: A (All guests are drunk), E (No guest is drunk), I (Some guests are drunk), O (Not all guests are drunk). Now, if in fact no guest is drunk, it is obvious that the best message is E, since after processing it the listener assigns 1 to the world where no guest is drunk (and the three other messages fail to achieve the same effect). Similarly if in fact all guests are drunk, the best message is A. The interesting case is the some-but-not-all world (denoted by $w_{\exists\neg\forall}$). In such a situation the two messages that could be used are I and O. Given (7), the speaker will choose I if and only if the following holds:

(8) $L(w_{\exists\neg\forall}|I) > L(w_{\exists\neg\forall}|O)$

Now:

(9) a. $L(w_{\exists\neg\forall}|I) = P_0(w_{\exists\neg\forall}|[\![I]\!]) = P_0(w_{\exists\neg\forall}|\{w_{\exists\neg\forall}, w_\forall\}) = \frac{P_0(w_{\exists\neg\forall})}{P_0(w_{\exists\neg\forall})+P_0(w_\forall)}$

b. $L(w_{\exists\neg\forall}|O) = P_0(w_{\exists\neg\forall}|[\![O]\!]) = P_0(w_{\exists\neg\forall}|\{w_{\exists\neg\forall}, w_{\neg\exists}\}) = \frac{P_0(w_{\exists\neg\forall})}{P_0(w_{\exists\neg\forall})+P_0(w_{\neg\exists})}$

It then follows straightforwardly that:

(10) I is better than O as a message *if and only if* $P_0(w_\forall) < P_0(w_{\neg\exists})$.

Equivalently (with $s_\exists = \{w_{\exists\neg\forall}, w_\forall\}$ and $s_{\neg\forall} = \{w_{\exists\neg\forall}, w_{\neg\exists}\}$):

(11) I is better than O as a message *if and only if* $P_0(s_{\neg\forall}) > P_0(s_\exists)$,
i.e., if and only if $P_0([\![O]\!]) > P_0([\![I]\!])$

This was the expected result: the speaker chooses the message which expresses the proposition that was the least likely to be true given the prior distribution, i.e, whose *surprisal value* is the highest.

Now, as noted by a reviewer, this prediction crucially relies on the view that the speaker, when choosing her message, measures its informativity (surprisal value) in terms of its literal interpretation. That is, we are discussing here the behavior of the fully rational version of the level-1 Speaker of the Rational Speech Act model, who assumes she is talking to a literal listener. A more sophisticated speaker might assume that she talks to a pragmatic listener. Because the pragmatic listener would interpret both 'some' and 'not all' as meaning 'some but not all', both are going to be equally informative, and our proposal in this paper could not work if we modeled the speaker in this way. We think we can argue for our choice on the following grounds. First, as discussed in the next section, it seems to be a fact that the choice between 'some' and 'not all' is partly governed by the relative infor-

always pick the best message, but, rather, picks each message with a probability which is increasing with the informativity of the message. In this particular respect, the model presented in this section is similar to the earlier Optimal Answer model of Benz and Van Rooij (2007) and to Franke's (2011) Iterated Best Response Model, among others. See Footnote 16 for further discussion.

[5] In principle, two messages could be exactly tied and be both optimal, so $S(w;L)$ is not a function. In our model, this will happen when the prior probability distribution P_0 of the listener is such that $P_0(s_\exists) = P_0(s_{\neg\forall})$ For simplicity, we ignore the possibility of such a tie, that is, we do as if such a prior distribution were not possible, which does not affect our conclusions in any way.

mativity (measured in terms of surprisal) of their literal meaning (cf. our discussion of the examples in (12) and (13) below). Second, scalar implicatures are in any case not always derived by the listener (in the experimental literature on scalar implicatures, rates of scalar implicature derivation are never very close to 100%, and are typically lower than for prototypical entailments). If the speaker believes that there is a small chance that she is talking to a literal listener (or a listener who believes that the speaker is not knowledgeable about the alternative, or one who takes the relevant alternative - say 'all' in the case of 'some' - to be irrelevant), she will always be better off choosing the message whose literal meaning is the most informative.[6]

5.2.2 Choosing between *I* and *O*

Is the prediction in (11) correct?

Imagine that we are talking about an international scientific conference where it is expected that everybody will give her talk in English. If the speaker happens to know that, contrary to expectations, some talks will not be given in English but in French, using (12a) below seems much more appropriate than using (12b).

(12) a. Not every talk will be in English.
b. Some talks will be in English.

Correspondingly, (13b) seems much more appropriate than (13a) in the same situation:

(13) a. Not every talk will be in French.
b. Some talks will be in French.

This is entirely in line with an explanation based on probabilistic expectations. In the specified context, the prior probability of the proposition that all talks will be given in English is higher than that of the proposition that no talk will be given in English. Correspondingly, the prior probability of (12a) (resp. (13b)) is smaller than that of (12b) (resp. (13a)), which predicts a preference for (12a) (resp. (13b)).

As observed by Roni Katzir (p.c.), one independent reason for the felicity contrasts in (13) and (12) might be due to the fact that *some* triggers a *not-many*-implicature, while *not every* might trigger a *many*-implicature. That is, for instance, (12b) might be interpreted as *some talks will be in English, but not many of them*, which in the situation we described might not be what the speaker intends to convey. And (13b) would be interpreted as conveying *some but not many of the talks will be in French*, which, on the contrary, might be exactly what the speaker wants to say. If the choice is between these two meanings, then only if the speaker happens to believe that many talks will be in French would she choose (13a), and this might play a role in the observed contrasts. However, we don't think the contrasts are markedly different if *some* is replaced with *many*. Suppose we are again in a context where it's highly expected (though not certain) that all talks will be in English, and we compare the following two sentences:

(14) a. Not every talk will be in English.
b. Many talks will be in English.

[6]We could of course consider a more sophisticated model where the speaker is uncertain as to whether she talks to the literal listener or to a pragmatic listener (or as to the value of other parameters, such as the underlying Question Under Discussion the listener is entertaining), but qualitatively the outcome would be the same.

It seems to us that in such a context, one would be much more surprised to hear (14b) than (14a). Yet if both sentences in (14a) are strengthened into 'many but not all talks will be in English', we should not observe such an effect. The contrast is, however, fully expected in our account, simply because (14a) expresses a proposition whose prior probability, in the specified context, is smaller than that of (14b).

5.2.3 Is *I* most often more informative than *O*?

Is it in fact the case that, in most contexts, the condition stated in (10) ($P_0(w_\forall) < P_0(w_{\neg\exists})$) holds? Even though this is very hard to assess based on actual data, there are good reasons to think that it holds, as already discussed in Chater and Oaksford (1999). These authors observe, first, that the properties denoted by nouns, verbs and adjectives typically hold of a minority of objects (they call this observation the 'rarity assumption'): there are less cats than non-cats, and less red things than non-red things and presumably most often there are less people who are singing than people who aren't. Note also that vague gradable adjectives like *tall* are typically interpreted in such a way that a minority of individuals within a comparison class count as tall (see Kennedy 2007 for discussion, among others). There are of course obvious counterexamples (*thing*, *exist*, ...), but overall, for most lexical predicates B, fewer things have the property B than the property non-B.[7] Second, Chater & Oaksford observed is that if predicates tend to be true of a small number of objects (as seems to be the case), then if we pick two predicates A and B randomly, we are much more likely to find that their intersection is empty than to find that A and B intersect (i.e., $P_0(s_\exists) < P_0(w_{\neg\exists})$), which entails the condition in (10).

In practice, however, the predicates *A* and *B* used in sentences of the form '*Q As* are *B*s' tend to be related in terms of their general subject matter. Typically, *A* denotes some small, cohesive region of the conceptual space (such as a species of animals, a nationality, a trade, the guests at a specific party, the talks at a specific conference, etc.), and *B* denotes a property that is well defined for *A*-objects (being of a certain color, doing a certain kind of activity, etc.). In these situations, the probability that No A is a B ($P_0(w_{\neg\exists})$) is not necessarily smaller than the probability that some As are B ($P_0(s_\exists)$). However, to the extent that, on most occasions, a randomly picked *A* is still more likely not to have property *B* than to have it (because predicates tend to denote minorities within a natural class), it will still be the case that it is less likely for all *A*s to have property *B* than it is for all *A*s not to have property *B*, which is exactly the condition stated in (10) ($P_0(w_{\neg\exists}) > P_0(w_\forall)$). Importantly, this condition is *weaker* than Chater and Oaksford's assumption that $P_0(s_\exists) < P_0(w_{\neg\exists})$, and is entailed by it. In Appendix 5.B, we provide a model which explains why a lexicon where lexical predicates *B* typically apply to a minority of objects within a natural class *A* is optimal from an information-theoretic point of view. From this reasoning, we expect *I* statements to be more informative than *O* statements in a majority of situations where both kinds of statements are true.[8]

[7] There are several reasons why this could be true. One is that 'natural' concepts typically cover a *connected* and relatively *homogeneous* region of the space of possible concepts (Gärdenfors 2004). To give an example, the *dog*-concept is arguably a more natural concept than the *non-dog* concept, because the concept of 'non-dog' includes many different types of objects which are intuitively extremely different from each other. This is the case even if we restrict our attention to a "natural" super-class of dogs, such as land animals or pets. In Appendix 5.B, we discuss a further reason why the 'rarity' assumption may hold, based on information-theoretic considerations.

[8] The fact that a *some*-statement is on average more informative than a *not-all*-statement is not sufficient by itself to explain why the former is lexicalized while the latter is not. Very informative messages are by

5.3 A model of the expected utility of a lexicon

The link between frequency of use and lexicalization can in principle be due to at least two types of pressures. One type of pressure is learnability: the more an expression is used, the easier it is to memorize it as a unit (see Hendrickson and Perfors 2019 for relevant discussion). The idea would be that, starting from a lexicon where the four corners are lexicalized, it will be easier for children to remember I than O, because they will hear I more often than O. As a result, such a language would be more likely to lose O than I when it is transmitted to the next generation. An explanation of this sort seems somewhat dubious in this specific case, since O-statements (expressed by 'not all' in English), though (as we argued) rarer than I-statements, are still not extremely rare. But there is in any case another well-known route to the same result (Zipf 1935; Piantadosi, Tily, and Gibson 2011, among many others): a language that lexicalizes frequent meanings as opposed to rare ones minimizes the average communicative effort of speakers (and parsing effort of listeners), compared to one where infrequent meanings, but not frequent meanings, would be lexicalized (in such a language, the meanings that you want to express most often require more words, and greater syntactic complexity). Conversely, speakers who seek to minimize their effort are likely to prefer inaccurate, simple expressions to accurate, complex ones; in a language where frequent meanings are lexicalized and therefore simple to express, this situation will be rarer and the average quality of information exchange will be higher.

In this section, we will offer a model of the *expected utility* of a lexicon in which a notion of cost is introduced. In any given situation, the model will assign a utility to each message, which will depend both on the cost of the message and its informativity (measured in relation with the prior probability distribution that characterizes the situation), and speakers are assumed to select the message that has the highest utility. We will be able to compute the *expected utility* of different lexica across occasions of uses. We will compare the expected utility of the attested lexicon $\{A, E, I\}$ with that of the unattested one, $\{A, E, O\}$. This approach formalizes the intuition that different lexicalization choices might lead to different outcomes in average communicative efficiency. In this section we offer a semi-formal account which contains the gist of our model (the fully explicit model is presented in Appendix 5.A).

We aim to compare the expected utility of two languages. In the first language, the quantifiers A, E and I are lexicalized, while the quantifier O is expressed by a syntactically complex expression, while in the second one, A, E, and O are lexicalized and I is expressed by a complex expression. We summarize this in (15), where the superscript $^+$ signals that a quantifier is more complex than all others, which will translate into a specific *cost* :

(15) a. $\mathcal{M}_I = \{A, E, I, O^+\}$,
b. $\mathcal{M}_O = \{A, E, O, I^+\}$.

We take the messages without the superscript to have a null cost, and the one with a superscript to have a positive cost c.

A situation of utterance consists of a pair $\langle w, P_0 \rangle$, where w is the actual world, by hypothesis known to the speaker, and P_0 is the probability distribution over worlds, corresponding to the beliefs of the listener in that situation (and that of the speaker before the speaker came to know w). As in RSA models that include message costs (Bergen, Levy, and Good-

definition true in fewer situations than less informative ones, so even though they are particularly useful when they are used, there are also fewer situations where they can be used. We certainly do not want to predict that in general messages corresponding to unlikely events or situations are more likely to be lexicalized!

man 2016), the utility of a message m in a situation $\langle w, P_0 \rangle$ is given by:

(16) $U(m, w, P_0) = \log(P_0(w|[\![m]\!])) - \text{cost}(m).$

Now, as discussed above, when w is either $w_\forall$ or $w_{\neg\exists}$, the message with the greatest utility is, in each language, A or E, since these messages have a null cost and are maximally informative. The comparison of the expected utility of each language thus hinges on what happens in the ∃¬∀-world. Let m be the 'cheap' message in the language under consideration (so $m = I$ in $\mathcal{M}_I$ and $m = O$ in $\mathcal{M}_O$) and m^+ be the expensive one (so $m^+ = O^+$ in $\mathcal{M}_I$ and $m^+ = I^+$ in $\mathcal{M}_O$).

We have:

(17) a. $U(m, w_{\exists\neg\forall}, P_0) = \log(P_0(w_{\exists\neg\forall}|[\![m]\!]))$
b. $U(m^+, w_{\exists\neg\forall}, P_0) = \log(P_0(w_{\exists\neg\forall}|[\![m^+]\!])) - c$

Now:

(18) $P_0(w_{\exists\neg\forall}|[\![m]\!]) = \frac{P_0(\{w_{\exists\neg\forall}\}\cap[\![m]\!])}{P_0([\![m]\!])} = \frac{P_0(w_{\exists\neg\forall})}{P_0([\![m]\!])}.$[9]

Likewise,

(19) $P_0(w_{\exists\neg\forall}|[\![m^+]\!]) = \frac{P_0(w_{\exists\neg\forall})}{P_0([\![m^+]\!])}$

Therefore,

(20) a. $U(m, w_{\exists\neg\forall}, P_0) = \log(P_0(w_{\exists\neg\forall}|[\![m]\!])) = \log \frac{P_0(w_{\exists\neg\forall})}{P_0([\![m]\!])}$
$= \log(P_0(w_{\exists\neg\forall})) - \log(P_0([\![m]\!]))$
b. $U(m^+, w_{\exists\neg\forall}, P_0) = \log(P_0(w_{\exists\neg\forall})) - \log(P_0([\![m^+]\!])) - c$

We assume that the speaker will pick the message with the greatest utility. That is, the speaker will use m^+ if and only if $U(m^+, w_{\exists\neg\forall}, P_0) > U(m, w_{\exists\neg\forall}, P_0)$, which is, given the equalities above, whenever the following condition holds:

(21) $\log(P_0([\![m]\!])) - \log(P_0([\![m^+]\!])) > c$

The *informativity* of a proposition ϕ relative to P_0, noted $\text{Info}_{P_0}(\phi)$ is defined (in information theory) as $-\log(P_0(\phi))$. That is, the less probable the proposition expressed by a message is, the more *surprising* and informative it is. So we can rephrase the above condition as:

(22) $\text{Info}_{P_0}([\![m^+]\!]) - \text{Info}_{P_0}([\![m]\!]) > c.$

This makes sense: it says that the speaker will choose the more costly message just in case the gain in information relative to the cheaper message exceeds the extra cost of the more costly message (in case she believes both messages).

When applied to each language, this gives us:

(23) a. In $\mathcal{M}_I$, the costly message (O^+), which means *not all*, is used (in a situation where the speaker believes the world is $w_{\exists\neg\forall}$) if:

$$\log(P_0(s_\exists)) - \log(P_0(s_{\neg\forall})) > c,$$

[9]Since m is either O or I, $\{w_{\exists\neg\forall}\} \cap [\![m]\!]$ reduces to $\{w_{\exists\neg\forall}\}$.

i.e., if:

$$\mathrm{Info}_{P_0}(s_{\neg\forall}) - \mathrm{Info}_{P_0}(s_{\exists}) > c;$$

otherwise the cheap message I is used.

b. In $\mathcal{M}_O$, the costly message (I^+), which means *some*, is used if:

$$\log(P_0(s_{\neg\forall})) - \log(P_0(s_{\exists})) > c,$$

i.e., if:

$$\mathrm{Info}_{P_0}(s_{\exists}) - \mathrm{Info}_{P_0}(s_{\neg\forall}) > c;$$

otherwise the cheap message O is used.

Now, in order to reason about expected utility, we need to take into account the fact that P_0 is not constant, that is, it varies across conditions of use and choices of predicates (technically, this means that P_0 is itself a random variable). We assume, following our discussion in Section 5.2.3, that while P_0 varies, it is more often the case that $P_0(s_{\exists}) < P_0(s_{\neg\forall})$ than the reverse (see Appendix 5.A for a precise way of expressing this assumption). Thus, the difference in informativity between I/I^+ and O/O^+, call it Q defined by (24), will have a probability distribution across situations that is biased towards positive values. Figure 5.1 illustrates what such a distribution might look like.

(24) $Q = \mathrm{Info}_{P_0}(s_{\exists}) - \mathrm{Info}_{P_0}(s_{\neg\forall})$

Focusing again on the case where the speaker believes $w_{\exists\neg\forall}$, we can essentially distinguish between two types of situations.

1. **Situations where the more costly message is used**.
   In both languages, the costly message is used in "some but not all" situations only if it is highly informative compared to the less costly one, so that the disadvantage it has in terms of cost is overridden. Now, in $\mathcal{M}_O$, where the costly message is I, this will happen when the prior probability of $s_{\exists}$ is sufficiently low compared to that of $s_{\neg\forall}$ (so that the costly message, which means $s_{\exists}$, will be highly informative), namely when $Q > c$. This corresponds to the right-hand region on Figure 5.1. Meanwhile, in $\mathcal{M}_I$, this will happen when the prior probability of $s_{\neg\forall}$ is sufficiently low (so that the costly message, which means $s_{\neg\forall}$, is highly informative): this corresponds to the left-hand region on Figure 5.1, where $Q < -c$. Given our assumption that the first situation happens more often than the second one, we will use more often the costly message in $\mathcal{M}_O$ than in $\mathcal{M}_I$ — on Figure 5.1, this corresponds to the fact that the right-hand region has a larger area under the curve than the left-hand region. This creates a disadvantage for $\mathcal{M}_O$ on the cost side.

2. **Situations where the less costly message is used**.
   When the prior P_0 is not sufficiently biased so as to make the costly message optimal, speakers always use the simpler message in "some but not all" cases (i.e., I in $\mathcal{M}_I$ and O in $\mathcal{M}_O$). This corresponds to the middle area in Figure 5.1, where $|Q| < c$ (the absolute difference in informativity between the two messages is smaller than their difference in cost). In this case, the comparison of the utilities achieved by the two languages hinges on how informative the simpler message is relative to P_0: when P_0 is such that $P_0(s_{\exists}) < P_0(s_{\neg\forall})$, the most informative message is I, and speakers of $\mathcal{M}_O$, who say O, incur a loss of utility. When the priors are such that $P_0(s_{\exists}) > P_0(s_{\neg\forall})$, the most informative message is O, and speakers of $\mathcal{M}_I$, who say I, incur a symmetric loss

of utility. Now, because of our assumption that the first situation ($P_0(s_\exists) < P_0(s_{\neg\forall})$) holds most of the time, which we assume remains true when one restricts oneself to less biased priors, the former situation will be more frequent than the latter; on Figure 5.1, this corresponds to the fact that the right-hand half of the middle area is larger than the left-hand half. Essentially, $\mathcal{M}_I$ will sometimes lead to diminished informativeness, but this will occur less often than equivalent losses under $\mathcal{M}_O$. Here again, there is an advantage for $\mathcal{M}_I$, this time on the informativity side.

This informal reasoning suggests that on average, the speaker of $\mathcal{M}_I$ will receive a higher utility than that of $\mathcal{M}_O$, both because she will use the costly message less often, and because when both types of speakers use the cheaper message, the speaker of $\mathcal{M}_I$ will be, most often, more informative than the speaker of $\mathcal{M}_O$.

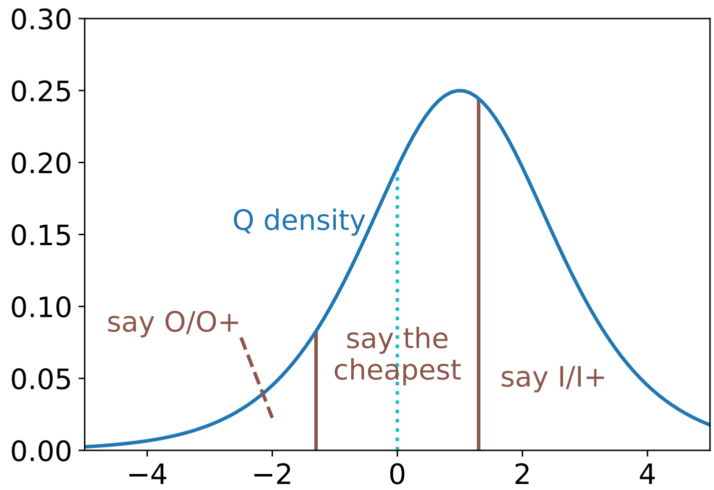

Figure 5.1 – An example of a density for $Q = \log(P_0(s_{\neg\forall})) - \log(P_0(s_\exists))$ ($= \mathrm{Info}_{P_0}(s_\exists) - \mathrm{Info}_{P_0}(s_{\neg\forall})$), biased towards larger values — which captures the fact that most often I-statements are more informative than O-statements. The x-dimension represents possible values of Q; it is divided into three intervals. In the first interval, speakers of both languages say O or O^+, because the meaning it expresses ($s_{\neg\forall}$) is much more informative than that expressed by I or I^+ ($s_\exists$); in the central interval, speakers of both languages say the cheaper message, because the difference in informativity (measured by Q) between the two messages is smaller than their difference in cost; in the third one, they say I or I^+, because its meaning ($s_\exists$) is much more informative than the meaning of O or O^+ ($s_{\neg\forall}$). The probability of each of these three cases is given by the area under the curve. The dotted blue line marks the limit between the domain where O, O^+ are more informative than I, I^+ and the one where the reverse is true.

In Appendix 5.A, we provide a formally explicit model which captures this reasoning.

5.4 Discussion

5.4.1 Summary

Let us take stock.

In Section 5.2 we suggested the following explanation for the non-lexicalization of O. English has a word for bakers but no word for people who don't sell bread, it has a word for dogs but no word for non-dog animals, and so on. It follows that, on average, an *O*-statement is less informative than an *I*-statement, so that in situations where both types of statements are true, speakers will most often use *I*. Given that less frequent meanings tend to be lexicalized less than frequent meanings, it is to be expected that *O* will not be lexicalized, across languages, to the same extent as *I* is.

In Section 5.3, we have offered a more explicit approach where we compare the overall expected utility of different lexica, showing that a lexicon based on $\{A, E, I\}$ has a higher expected utility than one based on $\{A, E, O\}$.

Our approach is limited in scope in that we only compare $\{A, E, I\}$ to $\{A, E, O\}$, but not to lexica with a different numbers of lexicalized corners of Artistotle's square. In particular, we do not really explain why lexicalized *O* is so rare. In principle, one might think that $\{A, E, I, O\}$ would be the ideal lexicon: speakers do feel the need to make *O* statements from time to time.[10]

We need to assume that some independent pressure to keep the lexicon minimal prevents the lexicalization of all four items. Ideally, this pressure would be part of our model.[11]

Similarly, one may wonder what is the expected utility of lexicalizing fewer corners. Thus, we would hope that lexicalizing $\{A, I\}$ is optimal within the two-element lexica, since this seems to be what is most common in natural languages (Katzir and Singh 2013). Taking compositionality into account (one can construct the missing messages by adding a negation), we would then compare the expected utilities of $\{A, E^+, I, O^+\}$, $\{A^+, E, I^+, O\}$ and $\{A, E, I^+, O^+\}$. It turns out that this time, there are terms of differing signs in the differences, and our current assumptions do not let us conclude as to the overall sign. Our approach therefore does not let us decide between these lexica, at least in its present form.

Finally, as we noted, the observation that *I* is more likely to be lexicalized than *O* holds not only in the domain of quantifiers over individuals, but also in the temporal and modal domains, as discussed by Horn (1973). Our approach would not have much difficulty to generalize to such cases, on the plausible assumption that, on average, the argument of such modal and temporal operators denote propositions which have a lower prior probability than their negation.

5.4.2 Comparison with other approaches

A strain of recent works have offered information-theoretic accounts for a number of properties of the lexicon of natural languages, based on the idea that languages maximize communicative efficiency through a trade-off between informativity and some notion of lexical complexity. This approach is most easily applied to content words, for which we can model the meaning space and the prior probabilities independently of linguistic facts. Examples include colour words (Zaslavsky et al. 2018), kinship terms (Kemp and Regier 2012), and animal names (Zaslavsky et al. 2019). The main challenge that any attempt to extend the idea

[10] In the model discussed in Section 5.3, they do so when *O*-statements are significantly more informative than *I*-statements, which will happen from time to time.

[11] One could argue that when one takes into account their scalar implicatures, *I* and *O* statements are truth-conditionally equivalent, in that they denote the $w_{\exists\neg\forall}$ situation. Therefore, there is no sense in lexicalizing operators for both; in fact, Aristotle's identification of four basic statements might be less relevant to natural language than the natural partition of possible worlds into just three sets. This is essentially the line of argument of Horn (1973, pp. 251–260). However, as we pointed out, *O* statements are attested and are not used in the same contexts as *I* statements (cf. our discussion of examples (12) and (13) in Section 5.2); for this reason, we are reluctant to just "shave off [the *O* category] with Occam's razor" (Horn 1973, p. 259).

to logical vocabulary faces is that it needs to define notions of informativity and complexity over the abstract domains that it discusses, and how to do so is not entirely straightforward.

Steinert-Threlkeld (2020), Denić, Steinert-Threlkeld, and Szymanik (2021), and Uegaki (2020) are among the works that take up this challenge. All three of these works show that the attested lexica of quantifiers (Steinert-Threlkeld 2020; Denić, Steinert-Threlkeld, and Szymanik 2021) or connectives (Uegaki 2020) in Natural Language tend to perform better than unattested lexica on a certain metric of efficiency, and Uegaki (2020) specifically points out that this fact can be seen as a solution to Horn's puzzle. They adopt specific notions of informativity and complexity for messages.

On the informativity side, some variant of expected listener surprisal, similar to what we ourselves use, is universally adopted. This requires a notion of prior probability on possible situations. Steinert-Threlkeld (2020) and Uegaki (2020) choose to represent situations as (some description of) possible worlds and to assign the same probability to each world. Thus, it is impossible for them to demonstrate any effect of "real-world" distributions. The flat prior also makes it so that informativity alone cannot distinguish upwards and downwards monotone quantifiers, in the sense that I and O (and A and E as well) will necessarily be equally informative, because they are true in exactly the same number of worlds. The solution to Horn's puzzle will therefore have to come from the complexity side. The way both Steinert-Threlkeld (2020) and Uegaki (2020) determine complexity is that they adopt a specific logical language based on common mathematical notation, including for instance the Boolean operators $\wedge$, $\vee$ and $\neg$, and they take a message's complexity to be the length of the shortest formula that represents it. As a consequence of this choice, downwards monotone operators are assumed from the get-go to be more complex than upwards monotone ones. We therefore expect lexica including O over I to be dispreferred, which is in fact what Uegaki (2020) finds. Thus, as a solution to Horn's puzzle, this approach is very similar to Horn's (1973) original proposal that there is an inherent semantic markedness to negation, and does not offer in turn an explanation for the latter fact (in this respect it is also similar to Katzir and Singh 2013). Where it improves on Horn's hypothesis is that it makes more specific predictions: Uegaki (2020) is able to explore the entire space of possible connective vocabulary.

An alternative to formulating arbitrary hypotheses about prior distributions and complexity is to derive them from linguistic data. This is what Denić, Steinert-Threlkeld, and Szymanik (2021) do. They adopt a classification of indefinites as well as a feature-based analysis due to Haspelmath (2001). They can then define the complexity of a message as the size of the smallest feature bundle that characterizes it. As far as the prior probability of each category is concerned, they estimate it from corpus data. Denić, Steinert-Threlkeld, and Szymanik (2021) do not discuss Horn's puzzle, and since they are exclusively concerned with indefinites, their model does not allow for a message meaning O.[12] We can still ask whether their method could offer a solution, if we were to adopt a similar classification of quantifiers (or connectives) and we found that inventories that satisfy Horn's generalization are more optimal. The main potential issue is one we already raised in Section 5.2. In principle, we want to derive informativity from the actual probability that a message is true. When we estimate the distribution of messages from corpus data, we are looking instead at the probability that a message is *produced*. If we consider that speakers take into account considerations of cost and complexity in their production, as in the model of Section 5.2.1,

[12] It should be noted that in saying that "languages lexicalize I," we have been abstracting away from the fact that most languages have a number of expressions with existential meaning, e.g. English *some*, *a*, *a certain*, *any*, *some … or other*, *whichever* etc.

then what we are measuring already reflects the effects of cost and of the vocabulary of the language. Thus, if we find that positive existentials (I) are more common than negated universals (O), it might be that this is because they are easier to express, and not the other way around. There is a parallel concern on the complexity side: if we base our representation of messages on morphological patterns or cross-linguistic lexicalization patterns, then we are going to assign a more complex representation to O from the beginning. These two biases will make it so that our model already internalizes Horn's observation and cannot be used as an explanation for it.

The conclusion of this discussion is that existing information-theoretic approaches to the logical vocabulary of languages would not offer a complete explanation for Horn's puzzle, because the models they use already internalize in some form either Horn's observation that O is uncommon, or Horn's hypothesis that negation is marked.[13] In contrast, we have derived the difference between the attested and the unattested lexicon entirely from the truth conditions of the messages, including a specific assumption about the prior probabilities of messages being true (as opposed to the probabilities of messages being used). We have also been able to derive the result analytically while leaving our assumptions somewhat abstract; for instance we have not been assuming particular probability distributions or particular costs. The price we pay is the extreme specifity of our result: as we have already noted, we are unable to extend the comparison beyond the two lexica that we discuss, and we do not account for any sort of pressure on lexicon size.

Our hope is that, despite the limitations we have just pointed out, and beyond the issue of Horn's puzzle, our work can serve as a further illustration of how explicit decision-theoretic models of pragmatics can in principle account, through the notion of expected utility, for certain universal tendencies in the logical lexicon of natural languages. Additionally, we hope to have shown that information-theoretic models can be used in linguistic research not just for broad, data-based analyses, but also to derive analytically specific qualitative points, in the same way as traditional formal analyses.

5.A Comparing the utilities of two lexica

Preliminaries In order to make the reasoning presented in Section 5.3 fully explicit, we need to make part of our model the fact that P_0, the prior distribution, is not given, and instead varies across situations. We can do this by making P_0 a random variable of its own, which ranges over possible probability distributions over the universe Ω (recall that Ω is the 3-element set $\{w_{\neg\exists}, w_{\exists\neg\forall}, w_{\forall}\}$).

This new assumption allows us to define the *expected utility* of a lexicon. For a lexicon $\mathcal{M}$ and a prior P_0, let us denote as $m(P_0, \mathcal{M})$ the message that a speaker of $\mathcal{M}$ will use in the situation $\langle w_{\exists\neg\forall}; P_0\rangle$ (when the world is $w_{\exists\neg\forall}$ and the prior is P_0). The utility achieved by

[13] In addition to the works we discussed, a somewhat different route is taken by Steinert-Threlkeld and Szymanik (2019), who show that certain semantic universals pertaining to quantifiers, such as permutation invariance, make quantifiers easier to learn by a Neural Network model. This could suggest that those universals emerge as an effect of learnability pressure. In this approach the learning properties of the network determine a notion of fitness that does not involve specific representational choices, beyond the architecture of the network itself. As in the other studies we discuss however, the observations that are presented to the model are drawn from an arbitrary distribution, so that no effect stemming from real-world prior distributions can be demonstrated; furthermore, formal properties of the model and the mode of presentation of the data (consisting in binary encodings of models) make it so that flipping downwards and upwards monotone quantifiers (such as flipping I and O, or A and E) cannot affect the results. Thus the model offers no insight towards Horn's puzzle.

the speaker in this situation can be written as $U(m(P_0, \mathcal{M}), w_{\exists\neg\forall}, P_0)$. Taking the expected value of this quantity over all potential prior distributions P_0, we obtain what we can call the conditional expected value of the lexicon,[14] which we write as $\bar{U}(\mathcal{M})$ and which is given in (25). This quantity represents the average utility achieved by the speaker of $\mathcal{M}$ in $w_{\exists\neg\forall}$-situations.

(25) $\bar{U}(\mathcal{M}) = \mathbb{E}[U(m(P_0, \mathcal{M}), w_{\exists\neg\forall}, P_0)]$

$\bar{U}(\mathcal{M})$ is what we will use to compare $\mathcal{M}_I$ and $\mathcal{M}_O$ in a formal way: what we want to derive is that $\bar{U}(\mathcal{M}_I) > \bar{U}(\mathcal{M}_O)$, corresponding to the idea that speakers of $\mathcal{M}_I$ achieve greater utility on average.

We also need to formalize the idea that P_0 is most often such that $P_0(w_{\neg\exists}) > P_0(w_\forall)$, that is, that in most situations speakers consider that no As being Bs is more likely than all As being Bs. There are probably various ways this could be done. Here is what we are going to assume: for any particular distribution P_0' which is biased in favor of $w_\forall$ relative to $w_{\neg\exists}$, we assume that the distribution P_0'' which encodes a bias of the same magnitude in the opposite direction is more likely to be the one that characterizes the listener's epistemic state. That is, for a particular distribution P_0', consider P_0'' which is like P_0', except that the probabilities of $w_{\neg\exists}$ and $w_\forall$ are flipped:

(26) $P_0''(w_{\neg\exists}) = P_0'(w_\forall)$
$P_0''(w_{\exists\neg\forall}) = P_0'(w_{\exists\neg\forall})$
$P_0''(w_\forall) = P_0'(w_{\neg\exists})$

If $P_0'(w_{\neg\exists}) > P_0'(w_\forall)$, then $P_0''(w_{\neg\exists}) < P_0''(w_\forall)$, and *vice-versa*. In other words, at most one of P_0' and P_0'' is such that the condition we expect to be the most common case is true. We are going to assume that this one is more likely to be the actual P_0 than the other:

(27) **Bias assumption (BA):** if Φ is the density of the variable P_0, and if P_0' and P_0'' are related in the way described above, then:
a. If $P_0'(w_{\neg\exists}) > P_0'(w_\forall)$, then $\Phi(P_0') > \Phi(P_0'')$.
b. If $P_0'(w_{\neg\exists}) < P_0'(w_\forall)$, then $\Phi(P_0') < \Phi(P_0'')$.

The assumption in (27) is how we will capture the fact that most of the time $P_0(w_{\neg\exists}) > P_0(w_\forall)$. If a particular choice of P_0 does not respect the condition, we assume that it is less likely than its mirror image that does respect it.

The BA is strictly stronger than our initial statement that "$P_0(w_{\neg\exists}) > P_0(w_\forall)$ is usually true", which one would most naturally implement as (28). (28) is in fact insufficient to derive the desired result: it might be that (28) is true, and that $\bar{U}(\mathcal{M}_I) < \bar{U}(\mathcal{M}_O)$. This will be the case for instance if the most probable values of P_0 are either such that $P_0(w_{\neg\exists})$ is much smaller than $P_0(w_\forall)$, or such that $P_0(w_{\neg\exists})$ is slightly greater than $P_0(w_\forall)$. However, we think that there is no reason why the distribution of P_0 should exhibit such an asymmetrical shape. The distributions commonly used in mathematical modeling usually have simple shapes that are obtained from smooth deformations of perfectly symmetrical ones,

[14] What makes $\bar{U}(\mathcal{M})$ "conditional" is that we only consider what happens in $w_{\exists\neg\forall}$-situations. It would be natural to define "expected utility" as the expectation of the utility achieved taken across all possible situations. $\bar{U}(\mathcal{M})$ is what we get if we condition on the fact that the world is $w_{\exists\neg\forall}$. The reason that we consider $\bar{U}(\mathcal{M})$ and not "proper" expected utility here is that the two languages we want to compare achieve exactly the same utility in $\neg\exists$-situations as well as $\forall$-situations, and therefore these situations will not matter to the comparison. In other words, whichever language yields a greater conditional expected utility also yields a greater expected utility.

with a single maximum or minimum. While real-world data can of course depart from this pattern, it is usually due to a well-identified categorical effect and we see no reason to think that it should happen in this instance. If we restrict ourselves to common parametric families, it is in fact the case that (28) and the BA are equivalent; in other words, any natural choice of parametrization for P_0 is such that the assumption in (28) would entail the BA.[15] For this reason, we do not think that our implementation of the bias assumption affects the generality of our result.

(28) $\mathbb{P}(P_0(w_{\neg\exists}) > P_0(w_{\forall})) > \frac{1}{2}$

The proof We can now prove the desired result: if the bias assumption holds, then $\bar{U}(\mathcal{M}_I) > \bar{U}(\mathcal{M}_O)$.

In Section 5.3, we derived the behaviour of speakers in our model. We can describe this behaviour in terms of the quantity Q defined in (29); such a description is given in (30).

(29) $Q := \log(P_0(s_{\neg\forall})) - \log(P_0(s_{\exists}))$

(30) a. If $Q < -c$, then $m(P_0, \mathcal{M}_I) = O^+$ and $m(P_0, \mathcal{M}_O) = O$.
b. If $Q > c$, then $m(P_0, \mathcal{M}_I) = I$ and $m(P_0, \mathcal{M}_O) = I^+$.
c. If $|Q| < c$, then $m(P_0, \mathcal{M}_I) = I$ and $m(P_0, \mathcal{M}_O) = O$.

Note that Q is exactly the difference in informativity between I and O: $Q = \mathrm{Info}(\llbracket I \rrbracket) - \mathrm{Info}(\llbracket O \rrbracket)$. This makes the above pattern intuitive: when I is much more informative, say I or I^+; when O is much more informative, say O or O^+; when the difference is small, say the cheapest. Since Q is a function of the random variable P_0, it is also a random variable.

What follows is a formal calculation that does not invoke any specific insight; the reader who is not interested in checking the correctness of it can skip straight to (36). We begin by decomposing $\bar{U}$ based on the value of Q:

(31)
$$\begin{aligned}\bar{U}(\mathcal{M}_I) = {} & \mathbb{P}(Q < -c)\mathbb{E}[U(O^+, w_{\exists\neg\forall}, P_0) \mid Q < -c] \\ & + \mathbb{P}(Q > c)\mathbb{E}[U(I, w_{\exists\neg\forall}, P_0) \mid Q > c] \\ & + \mathbb{P}(|Q| < c)\mathbb{E}[U(I, w_{\exists\neg\forall}, P_0) \mid |Q| < c]\end{aligned}$$

Furthermore, for any condition C:

[15]We think that the most natural choice of a parametrization for a 3-way probability distribution like P_0 would be a Dirichlet distribution $\mathrm{Dir}(\alpha, \beta, \gamma)$. Once we adopt this parameterization, (28) and the BA are both equivalent to $\alpha > \gamma$, so that the BA is innocuous.

Another, even simpler way to parameterize P_0 is to make the simplifying assumption that, when considering a sentence of the form QAB, the prior probability that a given A-individual x has property B is independent of the probability that some other individual A-individual y has property B, and that this probability is uniform across As. Then, P_0 depends entirely on the parameter p_0, the probability that a given A is a B, and the number of A-individuals. We have: $P_0(w_{\forall}) = p_0^n$ and $P_0(w_{\neg\exists}) = (1 - p_0)^n$, where n is the number of As. Thus the condition $P_0(w_{\neg\exists}) > P_0(w_{\forall})$ is equivalent to $p_0 < 0.5$, and (28) is equivalent to the density of p_0 having more mass on the left-hand side of the graph. If p_0 follows a Beta law (as would be most natural for a Bernoulli parameter), the distribution has a simple shape (e.g., a bell shape in the case where the parameters are greater than 1) and (28) will be true if and only if the density function tilts to the left, which would make the BA true as well.

A consequence of these facts is that if we had demonstrated our point through numerical simulations, as is common in the literature applying information-theoretic models to linguistics, the distributions we would have looked at would have been such that our result would have followed from (28). Because we want to derive an analytical result instead, we are forced to make our assumptions explicit, but all in all this makes our result more general, not less.

(32) $$\mathbb{E}[U(O^+, w_{\exists\neg\forall}, P_0) \mid C] = \mathbb{E}[\log P_0(w_{\exists\neg\forall}) - \log P_0(s_{\neg\forall}) - c \mid C] \\ = \mathbb{E}[\log P_0(w_{\exists\neg\forall}) \mid C] - \mathbb{E}[\log P_0(s_{\neg\forall}) \mid C] - c$$

And similarly:

(33) $$\mathbb{E}[U(I, w_{\exists\neg\forall}, P_0) \mid C] = \mathbb{E}[\log P_0(w_{\exists\neg\forall}) - \log P_0(s_{\exists}) \mid C] \\ = \mathbb{E}[\log P_0(w_{\exists\neg\forall}) \mid C] - \mathbb{E}[\log P_0(s_{\exists}) \mid C]$$

Putting these two together:

(34) $$\begin{aligned}\bar{U}(\mathcal{M}_I) = {} & \mathbb{P}(Q < -c)\mathbb{E}[\log P_0(w_{\exists\neg\forall}) \mid Q < -c] \\ & + \mathbb{P}(Q > c)\mathbb{E}[\log P_0(w_{\exists\neg\forall}) \mid Q > c] \\ & + \mathbb{P}(|Q| < c)\mathbb{E}[\log P_0(w_{\exists\neg\forall}) \mid |Q| < c] \\ & - \mathbb{P}(Q < -c)\mathbb{E}[\log P_0(s_{\neg\forall}) \mid Q > -c] \\ & - \mathbb{P}(Q > c)\mathbb{E}[\log P_0(s_{\exists}) \mid Q > c] \\ & - \mathbb{P}(|Q| < c)\mathbb{E}[\log P_0(s_{\exists}) \mid |Q| < c] \\ & - \mathbb{P}(Q < -c)c\end{aligned}$$

We can do the same thing with $\bar{U}(\mathcal{M}_O)$, and we derive:

(35) $$\begin{aligned}\bar{U}(\mathcal{M}_O) = {} & \mathbb{P}(Q < -c)\mathbb{E}[\log P_0(w_{\exists\neg\forall}) \mid Q < -c] \\ & + \mathbb{P}(Q > c)\mathbb{E}[\log P_0(w_{\exists\neg\forall}) \mid Q > c] \\ & + \mathbb{P}(|Q| < c)\mathbb{E}[\log P_0(w_{\exists\neg\forall}) \mid |Q| < c] \\ & - \mathbb{P}(Q < -c)\mathbb{E}[\log P_0(s_{\neg\forall}) \mid Q > -c] \\ & - \mathbb{P}(Q > c)\mathbb{E}[\log P_0(s_{\exists}) \mid Q > c] \\ & - \mathbb{P}(|Q| < c)\mathbb{E}[\log P_0(s_{\neg\forall}) \mid |Q| < c] \\ & - c\mathbb{P}(Q > c)\end{aligned}$$

When we take the difference, most terms cancel out:

(36) $$\bar{U}(\mathcal{M}_I) - \bar{U}(\mathcal{M}_O) = \overbrace{c(\mathbb{P}(Q > c) - \mathbb{P}(Q < -c))}^{\mathcal{C}} \\ + \underbrace{\mathbb{P}(|Q| < c)\mathbb{E}[\log P_0(s_{\neg\forall}) - \log P_0(s_{\exists}) \mid |Q| < c]}_{\mathcal{I} = \mathbb{E}[Q \mid |Q| < c]}$$

The remaining terms can be given intuitive interpretations. $\mathcal{C}$ is the difference in expected cost between the two languages. Speakers of $\mathcal{M}_I$ use costly messages when $Q < -c$, while speakers of $\mathcal{M}_O$ use costly messages when $Q > c$. The other term, $\mathcal{I}$, is the difference in expected informativity. The two languages result in different informativity only in situations where their speakers use the cheapest message, that is, when $|Q| < c$; in these situations, as we have seen, Q quantifies the difference in informativity.

In Section 5.3, we argued that both terms should be positive. This fact in fact follows from the bias assumption. To begin with, let us call ϕ the density of Q. It follows from the bias assumption that the following holds:

(37) For any $q > 0$, we have $\phi(q) > \phi(-q)$.

This is because mirroring P_0 as done per the bias assumption turns Q into $-Q$, and the variant that we assume to be more likely is also the one that yields a positive value for Q.

Then, we have:

(38)
$$\begin{aligned}\mathbb{P}(Q > c) - \mathbb{P}(Q < -c) &= \int_{c}^{+\infty} \phi(q)\mathrm{d}q - \int_{-\infty}^{-c} \phi(q)\mathrm{d}q \\ &> \int_{c}^{+\infty} \phi(-q)\mathrm{d}q - \int_{-\infty}^{-c} \phi(q)\mathrm{d}q \quad \text{(BA)} \\ &> \int_{-\infty}^{-c} \phi(q)\mathrm{d}q - \int_{-\infty}^{-c} \phi(q)\mathrm{d}q \\ &> 0.\end{aligned}$$

And:

(39)
$$\begin{aligned}\mathbb{E}[Q \mid |Q| < c] &= \frac{1}{\mathbb{P}(|Q| < c)} \int_{-c}^{c} \phi(q)q\mathrm{d}q \\ &= \frac{1}{\mathbb{P}(|Q| < c)} \left(\int_{-c}^{0} \phi(q)q\mathrm{d}q + \int_{0}^{c} \phi(q)q\mathrm{d}q \right) \\ &> \frac{1}{\mathbb{P}(|Q| < c)} \left(\int_{-c}^{0} \phi(q)q\mathrm{d}q + \int_{0}^{c} \phi(-q)q\mathrm{d}q \right) \quad \text{(BA)} \\ &> \frac{1}{\mathbb{P}(|Q| < c)} \left(\int_{-c}^{0} \phi(q)q\mathrm{d}q - \int_{-c}^{0} \phi(q)q\mathrm{d}q \right) \\ &> 0.\end{aligned}$$

It follows that $\bar{U}(\mathcal{M}_I) > \bar{U}(\mathcal{M}_O)$, as desired.[16]

5.B Expected utility and the optimality of predicates

We propose a model for a speaker who cares about classifying objects in a certain category A as being Bs or non-Bs. Let us take the word B to be defined only on As.[17] Having observed

[16]The result that $\bar{U}(\mathcal{M}_I) > \bar{U}(\mathcal{M}_O)$ still holds if, as suggested in Footnote 4, we make the assumption that speakers are only approximately rational, and that they pick their messages stochastically following a soft-max rule, as in the standard RSA model. While we do not provide a full proof, this follows from the following facts: (a) if P_0' and P_0'' are mirror images as in (26), then the expected utility achieved by speakers of $\mathcal{M}_I$ when $P_0 = P_0'$ is the same as the expected utility achieved by speakers of $\mathcal{M}_O$ when $P_0 = P_0''$, and vice-versa, and (b) if P_0 is such that $P_0(w_{\neg\exists}) > P_0(w_{\forall})$, then speakers of $\mathcal{M}_I$ achieve higher utility than speakers of $\mathcal{M}_O$ in the situation $\langle w_{\exists\neg\forall}; P_0 \rangle$.

However, once the assumption of total rationality is relaxed, we can no longer conclude from this technical result that $\mathcal{M}_I$ is the optimal language. Indeed, recall that $\bar{U}$ represents the *conditional* expected utility, as obtained if we only consider $w_{\exists\neg\forall}$-situations. With full rationality, we can ignore the other two situations ($w_{\forall}$ and $w_{\neg\exists}$), as speakers of both languages have the exact same strategy. Under the soft-max rule, however, this is no longer the case: speakers of either language can now use non-optimal messages in such situations (i.e., they can use I in $w_{\forall}$ and O in $w_{\neg\exists}$). Thus, we can no longer conclude that the comparison in terms of expected utility will go the same way as the comparison in terms of conditional expected utility. Hence, our proof does not generalize to a model where speakers are only approximately rational.

[17]This is a simplification. In a more realistic model, As and Bs would be subclasses of say, Cs. Assuming that B denotes a minority of Cs, if subclasses of C that get their own word are reasonably widely distributed over subsets of C, then B ought to also denote a minority of most of them. Thus the conclusion that most of the time, B ought to denote a subset of A doesn't crucially depend on the assumption that B is defined on As.

a new A, a speaker may want to tell others about it, and also whether it was a B. In this situation, the universe Ω is partitioned into two sets:

(40) NB: The A isn't a B.
B: The A is a B.

We ignore here the possibility of compositionally complex messages such as "A but not B".[18] Thus we assume there are two possible messages:

(41) A: "A!"
B: "B!"

Their semantics are the obvious ones:

(42) $[\![A]\!] = \Omega$
$[\![B]\!] = \text{B}$

We assume that they have the same cost, which allows us to simply ignore the cost term again. It is straightforward to verify that S_1 will always say B in a world in B, and A in a world in NB. We can then compute the *expected utility* of the message used by S_1 when she encounters an A and says something about it. This quantity expresses how useful their utterance is on average. We assume that the prior distribution over world-states is fixed and given by P_0. It represents both the actual probability that the A that S_1 encounters is a B, and the prior beliefs of the listener, who has not observed anything yet but has certain expectations before receiving information from S_1.

(43)
$$\begin{aligned}
\mathbb{E}_w[U_S(S(w;L_0)|w;L_0)] &= \sum_{w\in\text{B}} P_0(w)U_S(B|w;L_0) + \sum_{w\in\text{NB}} P_0(w)U_S(A|w;L_0) \\
&= \sum_{w\in\text{B}} P_0(w)\log\frac{P_0(w)}{P_0(B)} + \sum_{w\in\text{NB}} P_0(w)\log\frac{P_0(w)}{P_0(A)} \\
&= \overbrace{\sum_{w} P_0(w)\log P_0(w)}^{-H(P_0)} - \log P_0(B)\overbrace{\sum_{w\in\text{B}} P_0(w)}^{P_0(B)} \\
&\qquad - \log\underbrace{P_0(A)}_{1}\underbrace{\sum_{w\in\text{NB}} P_0(w)}_{1-P_0(B)} \\
&= -H(P_0) - P_0(B)\log P_0(B).
\end{aligned}$$

The first term does not depend on the lexicon: $H(P_0)$ depends solely on P_0, which is a parameter of the discourse context. However, the second term depends on what B means. In particular, imagine that speakers find themselves wanting to draw a new distinction within As, based on a specific binary feature. They could adopt a new word B^+, which refers to all As that have the feature one way, or B^-, referring to all As that have it the other way. This choice would have a consequence on the expected utility in (43), as the value of $P_0(B)$ wouldn't be the same: it could be either $P_0(B^+)$, or $P_0(B^-) = 1 - P_0(B^+)$. Then, we would expect speakers to actually adopt as B whichever term maximizes expected utility. We can verify that it is the rarer of the two: if $P_0(B^+) > 0.5$, it is optimal to choose B^- for B; if

[18] A more complete model could integrate complex messages, but it would make our calculations much more complex. A simple approach could be to include them, but assign to them a prohibitive cost.

$P_0(B^+) < 0.5$, it is optimal to choose B^+ for B. Thus, in the end, it is always optimal to have $P_0(B) < 0.5$. We therefore expect that a concept B is more likely to be lexicalized than its negation $\neg B$ if, for most natural categories A within which distinguishing Bs from non-Bs would be relevant, a random A-individual is typically less likely to be a B than a non-B. We do not predict through this reasoning that the prior of B should be particularly low, only that it should be below 0.5. Such a weak prediction is desirable given our discussion in Section 5.2.3.[19]

References

Barwise, Jon and Robin Cooper (1981). "Generalized quantifiers and natural language". *Linguistics and Philosophy* 4.2, pp. 159–259.

Benz, Anton and Robert Van Rooij (2007). "Optimal assertions, and what they implicate. A uniform game theoretic approach". *Topoi* 26.1, pp. 63–78.

Bergen, Leon, Roger Levy, and Noah D. Goodman (2016). "Pragmatic reasoning through semantic inference". *Semantics and Pragmatics* 9.

Buccola, Brian and Benjamin Spector (2016). "Modified numerals and maximality". *Linguistics and philosophy* 39.3, pp. 151–199.

Chater, Nick and Mike Oaksford (1999). "The probability heuristics model of syllogistic reasoning". *Cognitive psychology* 38.2, pp. 191–258.

Chemla, Emmanuel, Brian Buccola, and Isabelle Dautriche (2019). "Connecting Content and Logical Words". *Journal of Semantics* 3, pp. 531–547.

Denić, Milica, Shane Steinert-Threlkeld, and Jakub Szymanik (2021). "Complexity / informativeness trade-off in the domain of indefinite pronouns". In: *Semantics and Linguistic Theory (SALT) 30*, pp. 166–184.

Franke, Michael (2011). "Quantity implicatures, exhaustive interpretation, and rational conversation". *Semantics and Pragmatics* 4.1, pp. 1–82.

Gärdenfors, Peter (2004). *Conceptual spaces. The geometry of thought*. MIT press.

Geurts, Bart and Frans van der Slik (2005). "Monotonicity and processing load". *Journal of Semantics* 22.1, pp. 97–117.

Gibson, Edward et al. (2019). "How efficiency shapes human language". *Trends in Cognitive Sciences* 23.5, pp. 389–407.

Goodman, Noah D. and Andreas Stuhlmüller (2013). "Knowledge and implicature. Modeling language understanding as social cognition". *Topics in Cognitive Science* 5.1, pp. 173–184.

Haspelmath, Martin (2001). *Indefinite pronouns*. Oxford University Press.

Hendrickson, Andrew T. and Andrew Perfors (2019). "Cross-situational learning in a Zipfian environment". *Cognition* 189, pp. 11–22.

Horn, Laurence R. (1973). "On the semantic properties of logical operators in English". PhD thesis. University of California.

Katzir, Roni, Nur Lan, and Noa Peled (2020). "A note on the representation and learning of quantificational determiners". In: *Sinn und Bedeutung 24*. Vol. 1, pp. 392–410.

[19] Qing and Franke (2014) also use expected utility to explain why gradable adjectives such as *tall*, whose denotation seems to involve an unspecified threshold, still receive a non-trivial interpretation: their idea is that the threshold is expected by speakers to be set in such a way as to maximize expected utility. In fact, if one combines their idea with our definition of utility, it can be proven that a word like *tall* will optimally be true of a (large) minority of a given comparison class.

Katzir, Roni and Raj Singh (2013). “Constraints on the lexicalization of logical operators”. *Linguistics and Philosophy* 36.1, pp. 1–29.
Kemp, Charles and Terry Regier (2012). “Kinship categories across languages reflect general communicative principles”. *Science* 336.6084, pp. 1049–1054.
Kennedy, Christopher (2007). “Vagueness and grammar: The semantics of relative and absolute gradable adjectives”. *Linguistics and philosophy* 30.1, pp. 1–45.
Penka, Doris (2011). *Negative indefinites*. Oxford University Press.
Piantadosi, Steven T., Harry Tily, and Edward Gibson (2011). “Word lengths are optimized for efficient communication”. *Proceedings of the National Academy of Sciences* 108.9, pp. 3526–3529.
Qing, Ciyang and Michael Franke (2014). “Gradable adjectives, vagueness, and optimal language use. A speaker-oriented model”. In: *Semantics and Linguistic Theory (SALT) 24*, pp. 23–41.
Steinert-Threlkeld, Shane (2020). “Quantifiers in natural language optimize the simplicity/informativeness trade-off”. In: *22nd Amsterdam colloquium*, pp. 513–522.
Steinert-Threlkeld, Shane and Jakub Szymanik (2019). “Learnability and semantic universals”. *Semantics and Pragmatics* 12.
Uegaki, Wataru (2020). “*NAND and the communicative efficiency model”. Ms. University of Edinburgh.
Zaslavsky, Noga et al. (2018). “Efficient compression in color naming and its evolution”. *Proceedings of the National Academy of Sciences* 115.31, pp. 7937–7942.
Zaslavsky, Noga et al. (2019). “Semantic categories of artifacts and animals reflect efficient coding”. In: *41st Annual Conference of the Cognitive Science Society (CogSci)*.
Zeijlstra, Hedde (2011). “On the syntactically complex status of negative indefinites”. *The Journal of Comparative Germanic Linguistics* 14.2, pp. 111–138.
Zipf, George K. (1935). *The psychobiology of language*. Houghton Mifflin.

Chapter 6

Explaining presupposition projection in (coordinations of) polar questions

This chapter was originally published as:

> Émile Enguehard (2021). "Explaining presupposition projection in (coordinations of) polar questions". *Natural Language Semantics*.

Abstract

The chapter starts off with the observation that in certain cases, presuppositions triggered by an element inside a question nucleus may fail to project. In fact, in what looks like coordinated structures involving polar questions, presupposition projection patterns are exactly parallel to what is observed when the corresponding assertions are coordinated. It is further shown that these facts do not fall out straightforwardly from existing theories of polar questions, (apparent) coordinations of questions, and presupposition projection. I then propose a trivalent extension of inquisitive semantics such that the observed pattern can be understood in terms of existing theories of presupposition projection. The proposal has the following properties: (a) apparent coordinations of questions are indeed coordinations of questions, and (b) the semantic denotation of polar questions is asymmetric with respect to the "yes" and "no" answers.

Acknowledgements

My thanks go to the great number of people with whom I discussed part of all of this work and who provided me with helpful remarks, including in particular Benjamin Spector and two anonymous reviewers for *Natural Language Semantics*.

6.1 How presuppositions project from polar questions

6.1.1 Introduction: presupposition projection and coordination

The issue of presupposition projection in coordinated structures has been the subject of a significant amount of attention within the formal semantics and pragmatics literature. The basic pattern to be explained is known at least since Karttunen (1973) and goes as follows: utterance (1) below presupposes that Syldavia is a monarchy, which means among other things that (1) can be judged to be of degraded felicity in a context where the nature of the Syldavian political system is in doubt. The presupposition is due to the presence of a *trigger*, here the definite DP *the Syldavian monarch*. When the clause in (1) occurs in an embedded context, the presupposition may or may not "project", depending on certain factors. Of interest to us is the fact that the coordinated structures in (2) lack the presupposition that Syldavia is a monarchy.

(1) The Syldavian monarch is a progressive.

(2) a. Syldavia is a monarchy, and the Syldavian monarch is a progressive.
b. Syldavia is a republic, or the Syldavian monarch is a progressive.[1]

The lack of this presupposition for the examples in (2) is described and explained in various ways in the literature, going back in particular to the influential work of Karttunen (1973; 1974) and Heim (1983). At a minimum, an analysis of this data should derive the felicity conditions of (1) and (2) from more general assumptions about the semantics of declarative sentences, the semantics of the connectives *and* and *or*, and the semantics/pragmatics interface. For instance, Karttunen (1974) makes the following assumption about the semantics/pragmatics interface: a clause containing a trigger can be uttered felicitously if and only if its *local context* supports the trigger's presupposition. He further gives rules to determine the local context of conjuncts and disjuncts, allowing us to derive that the local context of the second conjunct in (2a) will support that Syldavia is a monarchy even when the global context does not, which in turn explains why the presupposition trigger *the Syldavian monarch* can be used felicitously in such a global context. Heim (1983) recasts Karttunen's (1974) analysis in a more general framework: she assumes that sentences denote *Context Change Potentials* (CCPs). In her model, one only has to specify how *and* and *or* string CCPs together to be able to derive the local context of every constituent clause.[2]

6.1.2 The core data

What I am interested in here is a very similar pattern of presupposition projection that can be observed in what looks like coordinations of *polar questions*, rather than coordinations of declarative clauses. Consider, to begin with, the simple polar question in (3a). As is well known, (3a) presupposes that Syldavia is a monarchy, in the sense that it can be infelicitous in a context where that fact is under doubt. A similar presupposition is observed when an

[1](2b) does presuppose that Syldavia is either a republic or a monarchy. I ask the reader to assume that this is a well-known fact about countries in the part of the world where Syldavia is located. The issue will not substantially affect the discussion.

[2]There is no explicit semantics for *and* in Heim 1983, and *or* is not discussed, but it is straightforward to apply Heim's approach to the discussion of *and* and *or* in Karttunen 1974 to reconstruct the full system.

embedded polar question occurs in a declarative sentence, as in (3b).[3]

(3) a. Is the Syldavian monarch a progressive?
b. Mary wonders whether the Syldavian monarch is a progressive.

The more general pattern is that a polar question, matrix or embedded, presupposes all that the corresponding declarative does — indeed, this fact is often used as a test to establish what declaratives presuppose.

English allows for what at least looks like conjunctions and disjunctions of polar questions.[4] We can investigate presupposition projection in such structures in much the same way as we do for declaratives. The result of this investigation is that presuppositions project from the second member in a coordination of questions following non-trivial patterns, and that the patterns in question are strikingly similar to what is observed for declaratives.

An example of what we will call a conjunctive question is given in (4). The apparent structure of (4) is $?p \wedge ?q$, where ? is a question-forming operator responsible for the auxiliary-fronting, p is *Syldavia is a monarchy*, and q is *the Syldavian monarch is a progressive*.[5] In spite of the fact that $?q$, the second conjunct of (4), appears to be the question in (3a), the speaker who asks (4) is not understood to be presupposing that Syldavia is a monarchy. In fact, in a context where it is established that Syldavia is a monarchy, it is infelicitous to ask (4), because (4) presupposes that whether Syldavia is a monarchy is still not known.[6]

(4) Is Syldavia a monarchy, and is the Syldavian monarch a progressive?

Thus, in (4), the presupposition triggered by the definite DP fails to project. In this respect, (4) is just like the conjunction of assertions (2a). As in (2a), the failure to project is due in some way or other to the presence of the first conjunct; if we replace it with something unrelated, as in (5), the presupposition does project. Again as for assertions, the utterance becomes infelicitous if the order is reversed to yield $?q \wedge ?p$, as in (6).That (6) is infelicitous in any context can be explained as follows: the presupposition projects, but the second conjunct somehow triggers an ignorance inference about p, and the two inferences are incompatible. This explanation is very similar to how one can explain why (7), $q \wedge p$, is odd: the presupposition projects, but then the second conjunct is necessarily trivial.

(5) Is Syldavia rich in mineral resources, and is the Syldavian monarch a progressive?

(6) # Is the Syldavian monarch a progressive, and is Syldavia a monarchy?

(7) # The Syldavian monarch is a progressive, and Syldavia is a monarchy.

More interestingly, we can try to replace the nucleus p of the first conjunct by a proposition p' that is (at least contextually) equivalent to $\neg p$, or to replace the first conjunct by an *or not* alternative question bearing upon the same proposition p. The reason these replacements are natural variants to test is that (8b) and (8c) intuitively raise the same issue

[3]What (3b) presupposes is arguably "Mary believes that Syldavia is a monarchy" rather than "Syldavia is a monarchy", or possibly both readings are possible. Whether the inferences I describe are attributed to the speaker or the attitude holder when they differ does not affect my discussion of them.

[4]Most connectives other than *and* and *or*, such as for instance *but*, do not seem to ever occur between two questions. Complex disjunctions such as *either... or...* cannot embed matrix questions either.

[5]We are going to discuss in time whether (4) should indeed be analysed as $?p \wedge ?q$.

[6]In general, matrix questions pragmatically presuppose that their answer has not yet been established in the discourse. Thus, the first conjunct in (4), uttered on its own, presupposes that Syldavia might or might not be a monarchy as far as the Common Ground allows. How this presupposition arises exactly in the case of (4) is an interesting question, but I will not attempt to address it here.

as (8a), in the sense that the knowledge one needs in order to answer any of these three questions is the same. However, they turn out not to be interchangeable when it comes to presupposition projection: both replacement attempts, (9a) and (9b), yield sentences that are just infelicitous.

(8) a. Is Syldavia a monarchy?
b. Is Syldavia a republic?
c. Is Syldavia a monarchy or not?

(9) a. # Is Syldavia a republic, and is the Syldavian monarch a progressive?
b. # Is Syldavia a monarchy or not, and is the Syldavian monarch a progressive?

The generalization that emerges is most easily phrased in terms of local contexts:

(10) **Generalization about conjunctive questions:** In a conjunction of polar questions $?p \wedge ?q$ (but not when *or not* questions are involved), while the local context of p is the global context, in at least some cases the local context of q is the global context enriched with p.

This generalization is entirely parallel to what Karttunen (1974) and Heim (1983) offer for the case of $p \wedge q$. It explains the felicity of (4), the infelicity of (6), the fact that the presupposition projects in (5),[7] and the infelicity of (9a) (the second conjunct's presupposition is necessarily not met).

Let us now turn to the case of disjunctive questions. An example of what looks like a disjunction of polar questions is given in (11); we can schematize it as $?p \vee ?q$, without committing to the idea it should be analysed that way. An immediate complication is that disjunctive questions of this sort are known to be systematically ambiguous between at least two readings. Using the terminology of Roelofsen and Farkas (2015), (11) has a *closed* reading where it presupposes (a) that one of John or Mary is here (exhaustiveness) and (b) that John and Mary are not both here (exclusivity). This reading is also known as the *alternative question* reading. The most natural way of bringing it out is to pronounce (11) with a falling intonation on the second disjunct. There is also an *open* reading, which is most easily made salient by having a rising intonation on the second disjunct. The open reading of (11) does not presuppose either (a) or (b), and is amenable to a negative answer ("Neither of them is here").

(11) Is John here or is Mary here?

With this distinction in mind, consider the crucial presuppositional example (12). The question in (12) can be uttered felicitously in a context where the political system of Syldavia is in doubt; in fact, as in the conjunctive case, (12) would be infelicitous otherwise.

(12) Is Syldavia a republic, or is the Syldavian monarch a progressive?

These facts obtain both with the "open" and "closed" intonation, and both under the open and closed reading. Note that we do not have to rely on the intonation to disambiguate readings: specifying a richer context lets us select one or the other reading and establish the presuppositional facts as well; for instance, the discourse in (13) brings about the closed

[7] To be precise, the fact that we infer from (5) that Syldavia is a monarchy, and not that if Syldavia is rich in mineral resources, it is a monarchy, is an instance of the *proviso problem*. The proviso problem occurs in exactly the same cases for questions as for assertions, as far as I can tell.

reading while the discourse in (14) brings about the open reading. In what follows, it will be helpful for the reader to keep these two discourses in mind to verify that all examples are compatible with both.

(13) My conviction is that you never see a conservative monarch enacting progressive policies. With what you told me about Syldavia's progressive laws, tell me, is Syldavia a republic, or is the Syldavian monarch a progressive? *(closed reading)*

(14) I always thought Syldavia was a very conservative monarchy, but what I learned about their policies made me less certain about it. Tell me, is Syldavia a republic, or is the Syldavian monarch a progressive? *(open reading)*

Thus, in (12), as in the disjunction of assertions (2b), the presupposition triggered by the definite DP fails to project. We can in fact replicate all of the tests we applied to the conjunctive case to see how the failure to project depends on the first disjunct. An unrelated first disjunct, as in (15), fails to block projection. Reversing the order as in (16) leads to degraded felicity (the judgment is less sharp than for the conjunctive case, and in this respect questions do not differ from assertions; cf. (17)). Replacing the first disjunct by its "opposite" or adding *or not*, as in (18a) and (18b) respectively, leads to infelicity.

(15) Is Syldavia rich in mineral resources, or is the Syldavian monarch a progressive?

(16) ?? Is the Syldavian monarch a progressive, or is Syldavia a republic?

(17) ?? The Syldavian monarch is a progressive or Syldavia is a republic.

(18) a. # Is Syldavia a monarchy, or is the Syldavian monarch a progressive?
b. # Is Syldavia a republic or not, or is the Syldavian monarch a progressive?

All of these observations lead us to the following generalization, which is strikingly similar to Karttunen's (1974) generalization about disjunction in declaratives:

(19) **Generalization about disjunctive questions:** In a disjunction of polar questions $?p \vee ?q$ (but not when *or not* questions are involved), under either an open or a closed reading, while the local context of p is the global context, in at least some cases the local context of q is the global context enriched with $\neg p$.

Thus, to conclude, presupposition projection in both (apparent) conjunctions and (apparent) disjunctions of polar questions obeys laws that are strikingly similar to those observed in conjunctions and disjunctions of declarative clauses. Yet as we are going to see, this pattern is puzzling from a theoretical point of view.[8]

[8]The availability of filtering in conjunctive questions has not been noted before to my knowledge, but Groenendijk (1998) (cited by Dotlačil and Roelofsen (2019)) mentions the example in (i), which demonstrates a parallel pattern when it comes to anaphoric dependencies. Both Groenendijk (1998) and Dotlačil and Roelofsen (2019) propose dynamic theories under which this example cannot be accounted for.

(i) Did you see a man? and was he angry?

Meanwhile, in a recent article, Abenina-Adar and Sharvit (2021) observe that filtering is available in disjunctive questions in certain cases similar to those we discuss here. They propose an account of disjunction based on alternative semantics where the observed projection pattern is specified by the semantics of the connective. We are going to adopt a very different approach here, among other things because we seek to understand conjunctive and disjunctive questions in parallel.

6.1.3 Some additional empirical points

A few other remarks can be made before we move on to the theoretical part of the paper. First, while we have been using matrix questions as examples, everything works exactly the same when looking at embedded questions; this is demonstrated below. In the disjunctive cases, the same ambiguity between open and closed reading is found in embedded examples, and as before it does not affect the projection facts we are interested in. In the rest of this article, we will switch between matrix and embedded examples depending on what makes a given point clearer, but as far as I can ascertain there is nothing in the discussion that does not apply equally to both.

(20) Mary wonders whether Syldavia is a monarchy, and whether the Syldavian monarch is a progressive.
$\not\rightsquigarrow$ (Mary believes that) Syldavia is a monarchy.

(21) Mary wonders whether Syldavia is rich in mineral resources, and whether the Syldavian monarch is a progressive.
$\rightsquigarrow$ (Mary believes that) Syldavia is a monarchy.

(22)
a. #Mary wonders whether the Syldavian monarch is a progressive, and whether Syldavia is a monarchy.
b. #Mary wonders whether Syldavia is a republic, and whether the Syldavian monarch is a progressive.
c. #Mary wonders whether Syldavia is a monarchy or not, and whether the Syldavian monarch is a progressive.

(23) Mary wonders whether Syldavia is a republic, or whether the Syldavian monarch is a progressive.
$\not\rightsquigarrow$ (Mary believes that) Syldavia is a monarchy.

(24) Mary wonders whether Syldavia is rich in mineral resources, or whether the Syldavian monarch is a progressive.
$\rightsquigarrow$ (Mary believes that) Syldavia is a monarchy.

(25)
a. ?? Mary wonders whether the Syldavian monarch is a progressive, or whether Syldavia is a republic.
b. #Mary wonders whether Syldavia is a monarchy, or whether the Syldavian monarch is a progressive.
c. #Mary wonders whether Syldavia is a republic or not, or whether the Syldavian monarch is a progressive.

Second, alternative descriptions of the phenomenon that do not refer to local contexts would be possible. However, as pointed out by Schlenker (2009) (cf. also Singh 2007), when it comes to assertions, the generalization in terms of local contexts lets us predict certain cases of infelicity due to triviality effects, even in the absence of presuppositions. Such effects are also observed in polar questions. For instance, in (26a), the clause "Ann is in France" should always be in a local context where Ann is known to be in Paris, and therefore locally trivially true. Similarly in (26b), the clause "Ann is in London" should always be locally trivially false. We can therefore understand the infelicity of (26a) and (26b) in terms of our generalization.[9]

[9] An anonymous reviewer points out that under certain intonations, these two examples can be felicitous. Discourse markers like *actually*, *in fact* etc. can help bring out such readings.

Similar data can be observed in the case of disjunction, as seen in (27a) and (27b). For now, we are going to focus on the presupposition projection data; we will return to examples like (26a) and (26b) in Section 6.4.2.[10]

(26) a. #Is Ann in Paris, and is she in France?
b. #Is Ann in Paris, and is she in London?

(27) a. #Is Ann away from Paris, or is she in France?
b. #Is Ann away from Paris, or is she in London?

6.2 Theoretical consequences and challenges

The fact that presupposition projection in coordinations of polar questions is very reminiscent of presupposition projection in coordinations of declarative clauses might seem expected. In this section, I am going to argue that it is actually puzzling, given established theories on polar questions, coordination of questions, and presupposition projection.

6.2.1 The need for a yes/no-asymmetry in polar questions

Some remarks can be made before we even attempt a formal analysis. To begin with, say we pursue an analysis where the following two properties hold:

(i) in a conjunctive question like (4) ($?p \wedge ?q$), there is actually a constituent identical or equivalent to the simple polar question (8a) ($?p$), as the syntax suggests;

(ii) whether a presupposition triggered in a certain context projects or not is a function of the semantics of the sentence and of its subconstituents (a property true of all analyses of presupposition projection we will discuss).

If so, then, we absolutely need our analysis of the simple polar question $?p$ to be *asymmetric*. What I mean by "asymmetric" here is that the semantic denotation of a polar question $?p$ should be such that the "yes" answer (p) and the "no" answer ($\neg p$) do not play interchangeable roles in it. In particular, $?p$ and $?(\neg p)$ should be different objects, and all three questions in (28) should be semantically distinct. The reason this is a necessity is that without a difference between these questions, there is no hope to account for the contrast between (4), on the one hand, and (9a) and (9b), on the other hand (a similar point can be made with the disjunctive examples).

(28) a. Is Syldavia a monarchy?

(i) Is Ann in Paris? And (actually) is she (even) in France?

Intuitively, in these cases, the speaker changes discourse strategy between the two conjuncts, and they would have asked the second question first, had they thought it through. The fact that the second conjunct is a correction is signalled through discourse markers and emphatic intonation. There is also a clear sentence break between the conjuncts.

The judgments reported here are based on a different intonation pattern, where there is no full sentence break between the clauses (which I have tried to indicate through punctuation) and the second conjunct does not bear any emphasis, and where there is no perception that the speaker changed their mind. Attempting an analysis of discourse-level uses of question coordination is beyond the scope of this paper.

[10] In particular, we will see that under the theory we are going to propose, it will be possible to explain the infelicity of (26a) and (26b) in terms of global redundancy or similar pragmatic constraints, without necessarily referring to local contexts.

b. Is Syldavia a republic?
c. Is Syldavia a monarchy or not?

The significance of this point comes from the fact that polar questions are frequently assigned denotations that are "yes/no-symmetric", e.g.:

(29) a. in Hamblin semantics: $\{p, \neg p\}$;
b. in partition semantics: $\lambda w. \lambda w'. p(w) = p(w')$;
c. in inquisitive semantics: $\{s \mid s \vdash p \vee s \vdash \neg p\}$.

Such theories make (28c) indistinguishable from (28a), and make (28b) contextually equivalent to the other two, which is problematic given that these are not interchangeable as far as the phenomenon we are looking at is concerned. We therefore need a more fine-grained view of polar questions.

It is important to mention that the fact that there exist yes/no-asymmetries around polar questions is well known, at least since Bolinger (1978). Some specific asymmetric phenomena that have been studied include discourse patterns, both in terms of felicity conditions (e.g. Biezma and Rawlins 2012) and future discourse potential (e.g. Roelofsen and Farkas 2015), as well as epistemic and evidential biases (e.g. Büring and Gunlogson 2000; Sudo 2013). These authors acknowledge the need for asymmetric representations of questions in order to account for the phenomena at hand. What is generally assumed, however, is that symmetric denotations of the kind shown above are adequate to account for the *resolution conditions* of polar questions, even if they might not be an adequate representation of question meaning in general. Under this view, what our data tells us is that presupposition projection in questions and/or the mechanisms of question coordination cannot be derived purely from resolution conditions.[11]

6.2.2 The failure of a direct reduction to the declarative case

A second remark has to do with the way we have to analyse the connectives. As already noted, the presupposition projection patterns we observe are very similar to those observed in the declarative case. Additionally, the basic facts we started from about how presuppositions project from declarative conjuncts and disjuncts are widely agreed-upon, and a large and diverse array of theories have been devised to derive them. If we can reduce our cases of "conjunctive questions" and "disjunctive questions" to conjunctions or disjunctions of declaratives, then said theories are going to make predictions about presupposition projection, which we might hope will match the observed pattern.

There is in fact an independent reason why we might want to reduce our examples to coordinations of assertions. The point would be to avoid committing to the fact that questions can be conjoined or disjoined at all. While the basic truth-conditional effect of *and* and *or* when acting upon declaratives is relatively well understood, there is much less of a consensus on what they do to questions. In various existing theories of questions, the meaning that should be assigned to *and*, to *or*, or to both to get non-degenerate results is unclear, or does not clearly relate to the declarative meaning, or there is a lack of uniformity between what has to be assumed for *and* and for *or*. Because of these pervasive issues, it has been proposed that questions can be conjoined but not disjoined, or at least not directly (e.g. Szabolcsi 1997; though see Szabolcsi 2016 and Ciardelli, Groenendijk, and Roelofsen 2018, sec. 9.2.2 for counterpoints), or that they can be disjoined but not directly conjoined

[11] I thank a reviewer for stressing the importance of this distinction.

(e.g. Hirsch 2017), or that they can be neither directly conjoined nor directly disjoined (e.g. Krifka 2001).

In these accounts of question coordination, which I am going to refer to collectively as the *reductionist view*, an apparent coordination of questions can always be paraphrased as a coordination of declaratives. For instance, one would analyse (30a) in such a way that it is equivalent to either (30b) or (30c), and (31a) in such a way that it is equivalent to (31b) or (31c).[12]

(30) a. Mary wonders whether John is here and whether it's raining.
b. Mary wonders whether John is here, and she wonders whether it's raining.
c. Mary wonders whether it is the case that John is here and it's raining.

(31) a. Mary wonders whether John is here or whether it's raining.
b. Mary wonders whether John is here, or she wonders whether it's raining.
c. Mary wonders whether it is the case that John is here or it's raining.

Unfortunately, such approaches do not lead to a satisfactory account. The low-scope strategy spelt out by (30c) and (31c) would make the presupposition projection facts unremarkable: the question nucleus would be a conjunctive or disjunctive proposition that is well known to lack a presupposition, and making a question out of this proposition should not change this. However, the problem is that our examples are not actually polar questions, or at the very least, some of the readings they allow for are not polar, and yet the presuppositions still fail to project. If (4) were a polar question, then the asker should be satisfied upon being told by a fully knowledgeable answerer that Syldavia's leader is a conservative, without being told whether we are talking about a monarch or a president. This is wrong: someone who asks (4) wants to actually know whether Syldavia is a monarchy or a republic. Similarly for the disjunctive case (12): we expect the asker to be satisfied upon being told that Syldavia cannot possibly have a conservative monarch. Under the open reading, such an answer is actually judged to be incomplete, and only felicitous if it is all the answerer knows. Under the closed reading, the answer is completely infelicitous, which is fully expected given that the asker is already presupposing the fact in question. Thus, if either of our examples has a polar reading at all, it is not very salient, and the lack of presupposition projection that we observe is not dependent on it.[13]

[12] Two strategies are possible. One option consists in assuming that the connective takes higher or lower scope than the surface syntax would suggest, with ellipsis or semantically inert elements potentially involved. For instance, Hirsch (2017) proposes that apparent conjunctive questions actually involve "and" taking high scope. The second option, proposed among others by Krifka (2001), consists in *lifting* the question from whatever type α our theory of simple questions would assign to it to the higher-order type $(\alpha \to t) \to t$, as shown in (i), and then assuming that conjunction and disjunction apply (classically) to this higher-order denotation. The second strategy only delivers "high" readings, as in (30b) and (31b).

(i) $\textsc{lift}(Q) := \lambda P.\, P(Q)$ (type $(\alpha \to t) \to t$)
"Q and Q'" $\approx \textsc{lift}(Q) \wedge \textsc{lift}(Q') = \lambda P.\, P(Q) \wedge P(Q')$
"Q or Q'" $\approx \textsc{lift}(Q) \vee \textsc{lift}(Q') = \lambda P.\, P(Q) \vee P(Q')$

To apply either strategy to the case of matrix questions, we need to assume that there are silent truth-conditional speech act operators above them, such as a silent *I wonder* or a silent *you should tell me*. Alternatively, along with the second strategy, we might assume that the higher-order type is the "normal" type of questions, and that there are no linguistic constituents with semantic type α. I use embedded examples in this section to avoid dealing with the issue.

[13] Another reason to think that our questions are not polar questions is that "yes" and "no" are not good answers to them. My impression from an informal survey is that in the conjunctive case (4), speakers' initial intuition is to accept "yes" and "no" as answers, but they are not sure as to how to interpret either. In the

The high-scope strategy for conjunctive questions, as spelt out by (30b), has been argued to predict adequate truth conditions in general (Krifka 2001; Hirsch 2017). For disjunctive questions, it has been proposed as an analysis of certain cases of disjunctive constituent questions (e.g. Xiang 2021). However, as already pointed out by Hoeks and Roelofsen (2019), the resulting truth conditions are clearly wrong in the case of disjunctive polar questions: on its most natural reading, (31a) is not a wide-scope disjunction (it does not suggest speaker ignorance) and it does not indicate a disjunctive desire on Mary's part.[14] Regardless of the truth conditions, in both the conjunctive and disjunctive cases, the high-scope strategy makes incorrect predictions about presupposition projection. Indeed, various theories of presupposition projection agree that "ϕ and ψ", where ϕ and ψ are declarative sentences, is felicitous whenever the presuppositions of ϕ are satisfied, and it is sufficient to grant ϕ to satisfy the presuppositions of ψ (I am ignoring here any issues related to the so-called proviso problem, as they do not affect the discussion as far as I can tell). Applying this to our conjunctive example (20), we predict that it should be felicitous in a context where, if we come to know that Mary wonders whether Syldavia is a monarchy, then we can grant that Syldavia is a monarchy. Such a context is extremely odd, and we clearly do not need to accommodate it to accept (20). Similarly, in the disjunctive example, we predict the presupposition not to project as long as, if Mary does not wonder whether Syldavia is a republic, then she believes that Syldavia is a monarchy — which again is not a context that one needs to accomodate to accept (23).[15]

To conclude, the low-scope reductionist view delivers inadequate polar interpretations for the questions we are looking at. The high-scope reductionist view makes specific predictions about presupposition projection, as long as we accept well-established facts about presupposition projection in declaratives, but these predictions do not bring us any closer to an explanation of what we observe. All this suggests that our data is problematic for reductionist accounts of question coordination, and that explaining this data will require that we actually engage with question semantics and with the thorny issue of question coordination.[16]

disjunctive case (12), with the "open" intonation, "no" is a felicitous answer, meaning that both disjuncts are false, but "yes" is of unclear interpretation. With the alternative question intonation, both "yes" and "no" are unacceptable. More generally, "yes" and "no" as answers to an alternative question can sometimes be interpreted as "both are true" and "neither is true", respectively, but they tend to come across as joke answers.

[14] An analysis following Xiang (2021), where *wonder* is decomposed into *want* and *know*, and disjunction takes scope inbetween the two, would predict such a disjunctive desire: Mary wants it to be the case that either she knows whether John is here or she knows whether it is raining.

[15] Under the 'want' > 'or' > 'know' scope analysis, the condition for filtering would instead be something like: "According to Mary's beliefs, if she does not know whether Syldavia is a republic, then Syldavia is a monarchy", which is nonsensical.

[16] A potential reductionist counter-argument worth commenting upon is that if we attempt to explicitly spell out the semantics that a reductionist account would give to our examples, it is not entirely clear whether the presupposition projects in the resulting sentence. This is especially true in the conjunctive case, (i), which is only somewhat degraded in my judgment. In the disjunctive case, the sentence we obtain, (ii), is somewhat involved and hard to interpret.

(i) ? Mary wonders / wants to know whether Syldavia is a monarchy and she wonders / wants to know whether the Syldavian monarch is a progressive.

(ii) ?? Mary wants to know whether Syldavia is a republic or to know whether the Syldavian monarch is a progressive.

A counter-counter-argument is that using complex conjunctions and disjunctions to make the scope explicit, even though it does not affect presupposition projection in the general case, makes both our attempts

6.2.3 Taking a broader look: our desiderata

In the previous two sections, I developed an argument that an analysis of our data should have the following two properties: it should feature a yes/no-asymmetric representation to polar questions, and it should not adopt a reductionist view of question coordination.

Some other properties are arguably desired in any account of question coordination. The effect of *and* and *or* when acting upon questions should relate in a natural way to their effect when acting upon propositions. Ideally, said relation should be the same for conjunction and for disjunction. The syntax we assume should be what we observe: two polar questions, each with auxiliary inversion or in embedded cases with *whether*, connected by *and* or *or*; in other words, $?p\wedge?q$ and $?p\vee?q$. Deviating from these requirements means that we need to assume a more complicated syntax/semantics interface. It might of course turn out that we need to violate some of them, but such a move requires good motivation.

Additionally, we want to account for the actual meaning of conjunctive and disjunctive questions (both open and closed), in terms of answerhood conditions, behaviour under embedding, and so on. This goes both as far as our specific examples are concerned and in the general case. The effect of *or not* should be explained as well.

The desiderata we listed so far bear entirely on the analysis of questions and of the question connectives that we adopt, and do not relate to the matter of presupposition projection directly. As I already mentioned, presupposition projection in declaratives is the subject of a large number of theories in the literature. These accounts have the following structure: they provide a system where the presuppositions of a complex declarative sentence can be derived from those of simpler sentences based on a few (ideally motivated) assumptions about the semantics/pragmatics interface and the denotation of sentences. Our goal here is to apply this methodology to questions. As we have seen above, we cannot get away with a reduction of our data to the sort of data that these theories have been designed to explain. However, we can still attempt to extend them in the most natural way. Concretely, what we want is to adopt the assumptions of an established theory of presupposition projection, apply them to an account of polar questions and coordinated questions, and derive the observed projection patterns.

How to do this, for a particular theory of presupposition projection and a particular theory of question semantics, is not necessarily obvious. In particular, it is difficult to extend a dynamic theory of presupposition projection, such as that of Heim (1983), to a static account of questions. In what follows, I am instead going to focus on theories of presupposition projection that are based on a static system, such as Schlenker's (2008) Transparency Theory (and the related derivation of local contexts he offers in Schlenker 2009), and what I am going to refer to as the *trivalent theory* (Beaver and Krahmer 2001; George 2014). The fact that these theories are based on a static and truth-conditional view of propositions will make it easier to extend them to various theories of questions.

My purpose in the rest of this section is to highlight the pervasive issues that one en-

---

more clearly contradictory:

(iii) a. Not only is Syldavia a monarchy, but the Syldavian monarch is a progressive.
b. #Not only does Mary wonder whether Syldavia is a monarchy, but she wonders whether the Syldavian monarch is a progressive.

(iv) a. Either Syldavia is a republic or the Syldavian monarch is a progressive.
b. #Mary wants either to know whether Syldavia is a republic or to know whether the Syldavian monarch is a progressive.

counters when applying the methodology I described to various established accounts of questions. The end goal is to motivate a new approach to question semantics, which will be detailed in the next section. We will proceed in the following way: first, in Section 6.2.4, we will review in detail how one could try to understand the phenomenon at hand within the framework of Hamblin/Karttunen (H/K) semantics. Then, in Section 6.2.5, I will briefly show how parallel issues to those that arise for H/K semantics arise in other major theories as well. This will motivate the presentation of a new theory of questions in Section 6.3.

6.2.4 Case study: an analytical attempt based on answer set semantics and the Transparency Theory

Probably the most common approach in question semantics is to analyse questions as denoting a set of answers. This approach is known as the *answer set theory* or as *Hamblin-Karttunen semantics*, after Hamblin (1976) and Karttunen (1977). I will now attempt an analysis of our data within H/K semantics, in order to show the many issues that arise.

The basic idea of H/K semantics is that a question is at some level a set of propositions, and these propositions are construed as the possible answers to the question. Asking a question means prompting the participants in the conversation to identify one of the propositions as true. There are various implementations of the idea. In a thread of work based on Hamblin's (1976) system, a question's denotation is simply the set of all its possible answers, regardless of the present situation. In contrast, under the approach of Karttunen (1977), questions have an intension and an extension, and the extension is the set of true answers at the current world. We are going to adopt Hamblin's view here for the sake of concreteness, but the choice does not really matter: the two systems map onto each other in a way that preserves the properties of interest to us (a formalization of this is offered in the Appendix, where both Hamblin's and Karttunen's systems are considered).

Presupposition projection In what follows, I will assume that presupposition projection is governed by the Transparency Theory for the sake of concreteness. My claim is that this choice is not crucial. The discussion can easily be recast in terms of local contexts or in trivalent terms, and in the Appendix I derive several formal results that illustrate the points made in the discussion in both the Transparency Theory and George's (2014) trivalent theory. Moreover, at least certain dynamic accounts of presupposition projection would face very similar issues when being extended to questions.[17]

I will not describe the Transparency Theory in full detail here, and only say how it applies to the examples: the Appendix contains proofs of some of the relevant formal results, as well as parallel results for trivalent theories of presupposition projection. The Transparency Theory's predictions derive from a notion of contextual equivalence over sentence denotations. Thus, to apply this theory to questions, we need a notion of contextual equivalence over the domain of questions. The simplest possible one goes like this: two questions are equivalent if the sets that they denote are equal after each member proposition is contextually restricted. We can use this simple notion to derive the results discussed here; cf. the Appendix for a formal definition (under the name H-equivalence) as well as potential alternatives.

[17] In particular, an account that combines the answer set theory with Heim's (1983) theory of propositions, like Li's (2019) "dynamicised Hamblin sets", would suffer the same problems as static answer set approaches.

General considerations In theories based on Hamblin 1976, a polar question like (32a) (repeated from (8a)) is traditionally analysed as denoting the two-element set in (32b). This analysis is based on the fact that the two propositions in (32b) are intuitively perceived to be the two ways one may felicitously answer (32a). More generally, one will analyse the polar question we schematize as $?p$ as denoting $\{p, \neg p\}$.

(32) a. Is Syldavia a monarchy?
b. {That Syldavia is a monarchy, That Syldavia is not a monarchy}

Sets being unordered objects, $?p = \{p, \neg p\}$ is indistinguishable from $?(\neg p)$ (at least under a classical view of negation where it is involutive). Thus at first look, the traditional view in H/K semantics fails to meet one of our desiderata: it does not analyse polar questions in a yes/no-asymmetric way. However, in some accounts of the internal syntax of disjunctive and polar questions, there is a yes/no-asymmetric intermediate constituent with the type of a question, which is the entity that is assumed to be engaging in disjunction. In particular, Biezma and Rawlins (2012) propose that a polar question $?p$ really denotes the singleton set $\{p\}$ rather than a two-element set, and that an additional "coercion operation" is responsible for adding $\neg p$ to the set. In a very similar way, Karttunen (1977) offers a syntax involving a "proto-question" constituent whose denotation is the equivalent of $\{p\}$ (translating from his system to Hamblin's). The negative proposition is added by a special semantic rule. In both theories, while the denotation of a plain polar question is yes/no-symmetric, question coordination can involve yes/no-asymmetric (proto-)questions, and thus our desideratum is met.

For purposes of illustration I am going to give the sketch of an analysis in this vein. Proto-questions are formed by a question operator ? such that $?p = \{p\}$. Then, a closure operator C, defined in (33), can apply. This operator generalizes the coercion operation of Biezma and Rawlins (2012) and the semantic rule of Karttunen (1977). What it does is add a "catch-all" answer to a question denotation — an answer that is true when all pre-existing answers are false. In particular, applied to $\{p\}$, it will give $\{p, \neg p\}$. The definition given here is sufficiently general that the operator C could in principle apply to non-singleton sets as well. Neither Karttunen (1977) nor Biezma and Rawlins (2012) allow for this possibility, but it will prove worth considering for our purposes.

(33) $$\mathrm{C} = \lambda Q.\, Q \cup \left\{ \neg \left(\bigvee_{p \in Q} p \right) \right\}$$

Following Karttunen (1977) as well as Biezma and Rawlins (2012), we assume that C is optional in at least some cases; in particular, we assume that while a plain polar question is essentially $\mathrm{C}[?p]$, a disjunctive question might be just $?p \vee ?q$.[18]

It would be most natural to assume that *and* and *or* denote set intersection and set union respectively when applied to questions. For disjunction, this is indeed the assumption that is generally made in H/K semantics. However, for conjunction, such an assumption leads to degenerate results. In most cases, the two questions being conjoined are disjoint sets; there

[18]The presentation here is not true to Karttunen (1977) and Biezma and Rawlins (2012) in the sense that neither theory actually features such an operator, and the corresponding device is more like an interpretation rule.

Additionally, Biezma and Rawlins (2012) assume that the corresponding coercion operation only occurs in embedded questions. Since the phenomenon we are interested in is not affected by the distinction between matrix and embedded questions, it will be easier to follow Karttunen (1977) and assume that both kinds are composed uniformly.

is no proposition that is an answer to both, and therefore their intersection is an empty set. We cannot hope to derive the felicity and answerhood conditions of conjunctive questions from the empty set. Because of that, one either needs to adopt a reductionist view of conjunction, or to assume that conjunction of questions applies *pointwise* inside the set. All this leaves us with the following definitions of question conjunction and disjunction:

(34) a. $Q \wedge Q' = \{p \wedge p' \mid p \in Q, p' \in Q'\}$
b. $Q \vee Q' = Q \cup Q'$

H/K semantics thus fails to fulfil another of our desiderata: the treatment of question conjunction and question disjunction is not uniform. This is disappointing in itself, inasmuch as it makes the syntax-semantics interface less transparent.[19] Additionally, it entails that we cannot hope to derive the presupposition projection patterns similarly for conjunction and disjunction, and therefore that the similarity between what is observed in questions and in declaratives will turn out to be a coincidence. This is an instance of a more general problem with H/K semantics, where the lack of parallelism makes the system not well-behaved from a formal perspective, leading to the impossibility of deriving formal properties and to strange predictions in edge cases (see Ciardelli, Roelofsen, and Theiler 2017 for a discussion of this point).

Conjunctive questions These concerns aside, let us move to an account of presupposition projection in coordinated questions. We have the following building blocks: polar questions denoting singleton sets, conjunction and disjunction as defined above, and the closure operator C.

The most common analysis of a conjunctive question with the apparent structure of $?p \wedge ?q$ in H/K semantics is that it should denote a 4-element set: $\{p \wedge q, p \wedge \neg q, \neg p \wedge q, \neg p \wedge \neg q\}$. I will call this set the *quadripartition*. The quadripartition leads to the prediction that the answerer to a conjunctive question has to specify the truth values of both propositions, resulting in four mutually exclusive answers. The received view is that this is indeed what conjunctive questions generally mean.

Recall that in the theory sketched above, $?p = \{p\}$, $?q = \{q\}$, and $\wedge$ applies pointwise. Additionally, C may apply on top of question-type constituents. We can therefore generate the quadripartition if we assume that C applies to both conjuncts separately, resulting in a structure like $\text{C}[?p] \wedge \text{C}[?q]$.

Yet, the quadripartition is problematic for us. To predict the filtering observed in (35) (repeated from (4)), we would want the proposition expressed by p (that Syldavia is a monarchy) to be transparent in the position of q (the second conjunct's nucleus). However, it is easy to verify that it is not; in fact, nothing non-trivial is transparent in the position of q in $\text{C}[?p] \wedge \text{C}[?q]$ in any context (cf. the Appendix for a proof). Regardless of the presupposition projection theory we adopt, we know at any rate that the quadripartition will not let us derive the observed pattern, because it is not yes/no-asymmetric: substituting $\neg p$ for p in the quadripartition does not change the set.

(35) Is Syldavia a monarchy, and is the Syldavian monarch a progressive?

[19] Here there are several things to say to defend H/K semantics. First, pointwise composition is independently motivated, not only to analyse the internal composition of questions (Hamblin 1976) but also for other phenomena, such as focus (Rooth 1992). Second, given that non-pointwise conjunction leads to degenerate results, it seems unproblematic to assume that the syntax allows for it just as it allows for non-pointwise disjunction, but that it is banned on semantic grounds.

Consequently, our data is problematic for the view that conjunctive questions always denote the quadripartition. It is also problematic for answer set theories not based on proto-questions, and where polar questions always denote $\{p, \neg p\}$, as in these theories the quadripartition is the only thing that a conjunctive question could denote. In our system, however, it is possible to generate other denotations, as long as we give ourselves as much freedom as we want and allow for C to apply to any question-type constituent. This allows us to generate six alternative denotations for (35), which are listed in (36).[20] This wide use of C is of course not at all in the spirit of Karttunen (1977) and Biezma and Rawlins (2012), for whom C is a last-resort coercion rule.

(36)
a. $?p \wedge ?q = \{p \wedge q\}$
b. $?p \wedge \mathrm{C}[?q] = \{p \wedge q, p \wedge \neg q\}$
c. $\mathrm{C}[?p] \wedge ?q = \{p \wedge q, \neg p \wedge q\}$
d. $\mathrm{C}[?p \wedge ?q] = \{p \wedge q, \neg(p \wedge q)\}$
e. $\mathrm{C}[?p \wedge \mathrm{C}[?q]] = \{p \wedge q, p \wedge \neg q, \neg p\}$
f. $\mathrm{C}[\mathrm{C}[?p] \wedge ?q] = \{p \wedge q, \neg p \wedge q, \neg q\}$

Options (a), (b), and (c) are uninteresting. (a) leads to a question with only one possible answer; such a question is presumably not well-formed. (b) and (c) should correspond to readings where the answer to one of the conjuncts is presupposed, which is presumably banned on pragmatic grounds.

Option (d) leads to a yes/no question. We have already seen when discussing the reductionist view that (35) is not, or at least does not have to be, a yes/no question.

Option (f) is yes/no-symmetric as well (with respect to the first conjunct), leaving us with option (e). Option (e) is in fact what we want. First, p is transparent in the position of q. Intuitively, this is because in the denotation, "q" always appears somewhere after "$p\wedge$".[21] We therefore correctly predict (4) to lack a presupposition that Syldavia is a monarchy.[22] Furthermore, the three mutually exclusive propositions in the set, which I will refer to as the *tripartition*, correspond exactly to the possible answers to (35):

(37)
a. Syldavia is not a monarchy. ($\neg p$)
b. Syldavia is a monarchy and/but the Syldavian monarch is not a progressive. ($p \wedge \neg q$)
c. Syldavia is a monarchy and the Syldavian monarch is a progressive. ($p \wedge q$)

Thus, what we now have is a new analysis for conjunctive questions within H/K semantics, schematized by (e), such that both the presupposition projection facts and the answerhood conditions of (35) are easily understood.[23]

In my view, the analysis in question is not entirely satisfactory. Other than the general problems with H/K semantics and pointwise conjunction I mentioned above, the main issue is with the asymmetric structure of (e). This asymmetry raises a number of challenges

[20] Adding extra C's would only add logical contradictions to the sets, as the reader can verify. We are going to assume that it would make the questions unacceptable.

[21] Since all theories of presupposition projection are designed so that a presupposition that p is true does not project after "$p\wedge$", here we can see that the choice of the Transparency Theory was not crucial. See also the derivations in Appendix 6.A.

[22] Our generalization in terms of local contexts given in (10) also follows if we derive local contexts in the spirit of the Transparency Theory, along the lines of Schlenker 2009.

[23] We will discuss in Section 6.3.5 the implications of the resolution conditions predicted by (e) in non-presuppositional examples.

if we make the natural assumption that C is in the syntax and that it is some sort of complementizer. First, assuming that coordination can only occur between constituents of the same category, we need to explain why a constituent like $?p \wedge \text{C}[?q]$ is at all possible. Second, we need to explain why (e) is allowed, whereas the alternative tripartition in (f), which is not attested, is not. Even the polar reading predicted by (d) might have to be ruled out, as the evidence for such a reading is scarce.[24] Finally, it is puzzling that nothing in the surface syntax of (35) suggests there is left/right asymmetry in the structure. This is probably why, to my knowledge, the possibility of something like (e) (or (f)) has in fact never been considered worth discussing before.[25] None of these problems is insurmontable, but they will require us to complicate the syntax, or else to find another analysis of the mechanism responsible for the operation performed by C where they become easier to deal with.

To conclude, while H/K semantics provides the necessary ingredients to understand our cases of conjunctive questions, doing so requires non-standard assumptions about conjunctive questions' internal composition (with unconstrained, asymmetric application of a closure operation), so as to allow them to be interpreted as a tripartition, and possibly suffers from an overgeneration of readings.

Disjunctive questions We can apply the same methodology to our disjunctive example (38) (repeated from (12)). The most common analysis of closed disjunctive questions (alternative questions) within H/K semantics is that a question with the apparent structure $?p \vee ?q$ denotes $\{p, q\}$, based on the intuition that the two potential answers to the question are the two disjuncts (cf. discussion by Roelofsen and Farkas (2015)). Both Karttunen (1977) and Biezma and Rawlins (2012) essentially adopt this view.

(38) Is Syldavia a republic, or is the Syldavian monarch a progressive?

Assuming that disjunction is set union, $\{p, q\}$ can easily be derived as the direct disjunction of the proto-questions $\{p\}$ and $\{q\}$, i.e. $?p \vee ?q$. This is in fact part of the motivation for proto-questions: if a polar question denoted $\{p, \neg p\}$ from the get-go, a disjunction of two such questions would be $\{p, \neg p, q, \neg q\}$. Yet there is no reading of disjunctive questions such that the set of felicitous answers is this 4-element set. For instance, a fully knowledgeable speaker cannot cooperatively answer "John is not here" ($\neg p$) to (39), whatever the intonation.

(39) Is John here, or is Mary here?

The version of H/K semantics we are working with here partially solves the problem by allowing for $?p \vee ?q = \{p, q\}$. We do generate $\text{C}[?p] \vee \text{C}[?q] = \{p, \neg p, q, \neg q\}$, but we can rule it out by assuming a *unique-answer constraint*: at every contextually permissible world, there should be exactly one answer that is true. It can be verified that there is no context where $\{p, \neg p, q, \neg q\}$ satisfies this constraint without making some of the answers trivially false in the context (see Hoeks and Roelofsen 2019 for extensive discussion of the interaction between disjunctive questions and various refinements of the unique-answer constraint).

[24] Hoeks and Roelofsen (2019) note that, like the theory we are sketching here, inquisitive semantics predicts conjunctive questions to have polar readings, and they claim that this prediction is correct. My own judgment is that the reading they characterize as polar is the one corresponding to the tripartition, that is, option (e). In my view, the issue remains to be investigated more thoroughly.

[25] In particular, the extensive discussion of potential structures for conjunctive questions offered by Hoeks and Roelofsen (2019) does not mention the possibility of (e) or (f).

The open reading of disjunctive questions has been the subject of less theorizing, as it is often mischaracterized as a polar reading. As Roelofsen and Farkas (2015) argue, the open reading is not polar and in fact allows for three possible answers corresponding to the disjuncts (p, q) and to their joint negation ($\neg p \wedge \neg q$). The natural analysis, therefore, is $\{p, q, \neg p \wedge \neg q\}$. We can generate it as $\textsc{c}[?p \vee ?q]$, as Roelofsen and Farkas (2015) propose (Roelofsen and Farkas (2015) work within the framework of inquisitive semantics, but the variant of H/K semantics that we are using here is close enough that the analysis carries over).[26]

We therefore have two potential denotations, $\{p, q\}$ and $\{p, q, \neg p \wedge \neg q\}$. Unfortunately for us, neither of them leads to an immediate explanation of the projection facts. Recall that in the conjunctive case, the key property that we could capitalize on, under the Transparency Theory but also under other theories of presupposition projection, is that q always appeared after "$p\wedge$" in the denotation. There is no similar property in either of our proposed denotations: q appears on its own.[27] This translates into a failure of transparency: for $\neg p$ to be transparent in the position of q in either case, a necessary condition is that $\{p, q\} = \{p, \neg p \wedge q\}$ for any arbitrary proposition q (cf. also Appendix 6.A). There is no reason for this to be the case in any realistic example.

At this point, one might remark that the notion of question equivalence that we have been using so far is naïve and excessively fine-grained. To fix this, first, we would have to define equivalence relative to a context, as Schlenker (2008) in fact does. Additionally, we might want to take the unique-answer presupposition into account. There are many natural approaches that one could think of to that effect — Schlenker's theory as laid out in Schlenker 2008 does not cover the case of multiple sources of presupposition in the same sentence — and in the Appendix I prove for a number of them that they do not let us derive the desired result. Thus (12)/(38) remain entirely mysterious from the perspective of H/K semantics and the Transparency Theory.

In the end, H/K semantics lets us explain presupposition projection in conjunctive questions under the Transparency Theory (among others), if we are ready to make certain *ad hoc* assumptions on question composition. More problematically, while there is a natural analysis for open and closed disjunctive questions within H/K semantics, the observed presupposition filtering does not follow in a similar, mostly theory-neutral way, and neither does engaging in detail with the Transparency Theory (as is done in the Appendix) let us derive it. At any rate, the fact that conjunctive and disjunctive questions receive a very different treatment and that considerations from one case do not extend straightforwardly to the other is unsatisfactory, given that we hope to explain the parallelism between questions and assertions. Thus, given the desiderata exposed above, our data is problematic for the H/K analysis of polar questions, even in its proto-question-based yes/no-asymmetric form.

6.2.5 Other theories in question semantics

We have seen that our data is puzzling from the perspective of the answer set theory of questions. We are naturally prompted to look for alternative views on question semantics that do not suffer from the same problems as H/K semantics. While I lack the space to

[26] In contrast, Biezma and Rawlins (2012) analyse both open and closed questions as $\{p, q\}$, but with different presuppositions.

[27] As in the conjunctive case, there are in principle other possibilities, such as $?p \vee \textsc{c}[?q] = \{p, q, \neg q\}$ and $\textsc{c}[?p] \vee ?q = \{p, \neg p, q\}$. These options do not have the required property either; I will ignore them here to keep the discussion contained.

review these views at the same level of detail, I will discuss the most common approaches here and show that largely the same issues arise.[28]

Questions as partitions Groenendijk and Stokhof (1984) propose (as a component of a more complex theory) an analysis of questions based on the following intuition: what questions contribute to the conversation is that they raise an *issue*, and the issue is a partition of the logical space. Thus, the denotation of a question will be a partition, represented in the form of an equivalence relation over possible worlds (type $s \to s \to t$). Two worlds are related if and only if they resolve the issue in the same way. Concretely, this will be the denotation of a polar question $?p$:

(40) $\lambda u.\, \lambda v.\, p(u) = p(v)$

The partition theory fails to meet our desiderata in a more fundamental way than H/K semantics. First, the denotation of a polar question is yes/no-symmetric: substituting $\neg p$ for p in (40) would keep the object unchanged. Unlike in the answer set theory, there is no obvious fix without deeply changing the account. A related fact is that a conjunction of two questions will denote what I have called earlier the quadripartition, where the desired transparency property does not obtain, and there is no obvious way to derive the tripartition that we want. Finally, since a disjunction of equivalence relations is not in general an equivalence relation, the partition theory predicts that disjunctions of questions will be ill-formed. It therefore needs to resort to reductionist techniques to analyse question disjunction; we have seen why this is problematic for our purposes.

Inquisitive semantics Inquisitive semantics (Ciardelli, Groenendijk, and Roelofsen 2013) is a framework in which questions and assertions are analysed as having the same type. Both are sets of sets of worlds, construed as the epistemic states that resolve the question or make the assertion true. Concretely, a proposition p is the set of sets of worlds such that p is uniformly true in the classical sense, as in (41a). Here "$s \vdash p$" is to be read as "s supports p" and means that all worlds in s are p-worlds. I will write P for the inquisitive denotation to distinguish it from the classical one, and I will use Q, Q', etc. for abstract variables denoting inquisitive propositions, even though that is not standard practice. A polar question over proposition P is assumed to have the structure $?P$, where $?$ is an operator defined in (41b), which adds to its argument all the sets that have no intersection with anything in it. The result is given in (41c). Notice that P is yes/no-asymmetric, but $?P$ is not.

(41)
a. $P = \{s \mid s \vdash p\}$
b. $? = \lambda Q.\, Q \cup \{s \mid \forall s' \in Q.\, s \cap s' = \varnothing\}$
c. $?P = \{s \mid s \vdash p\} \cup \{s \mid s \vdash \neg p\}$

The structure of the analysis is actually extremely similar to our version of H/K semantics. The basic denotation of a question/assertion is yes/no-asymmetric and derives directly from the corresponding proposition. Then, an additional operator adds the negative case. (42) gives an helpful translation table from H/K semantics to inquisitive semantics; notice

[28] All the theories discussed in what follows are static; while there exist also dynamic accounts of questions, I will set them aside for lack of space. In general, these accounts' handling of the connectives is derivative of a static theory, and the issues that we are going to discuss carry over. This applies for instance to Dotlačil and Roelofsen 2019 (based on inquisitive semantics) and to Li 2019 (based on answer set theory).

that the main innovation at this point is the loss of the distinction between propositions and proto-questions.

(42) $p \longrightarrow P$
$?p = \{p\} \longrightarrow P$
$\textsc{c}[?p] = \{p, \neg p\} \longrightarrow ?P$

Taking the maximal elements (under set inclusion) in an inquisitive proposition, called the *alternatives*, in fact lets us perform the opposite translation to H/K denotations. The operator ALT defined in (43) does just that.

(43) a. $\textsc{alt}(Q) := \{s \mid s \in Q \wedge \neg\exists s' \in Q.\, s \subsetneq s'\}$
b. $\textsc{alt}(P) = \{p\}$
c. $\textsc{alt}(?P) = \{p, \neg p\}$

Unlike H/K semantics, inquisitive semantics allows for parallel definitions of conjunction and disjunction as set intersection and set union respectively, which, as we already discussed, is desirable from a theoretical perspective. Relative to the H/K definitions, the inquisitive connectives do the same thing to alternatives in simple cases, and where they differ, the inquisitive behaviour has been argued to be preferable (Ciardelli, Roelofsen, and Theiler 2017). The definitions are given in (44).

(44) a. $Q \wedge Q' = Q \cap Q'$
b. $Q \vee Q' = Q \cup Q'$

Unfortunately, even though I made such uniform definitions part of our desidarata, in this case the uniformity does not bring us any closer to understanding presupposition projection. In fact, the discussion of H/K semantics above carries over almost entirely to inquisitive semantics, translating the structures as per (42) and using ALT to recover the predicted answers. For conjunctive questions, $?P\wedge?Q$ is the quadripartition, and we need $?(P\wedge?Q)$ to get the desired tripartition and transparency to obtain. The treatment of disjunctive questions is again not uniform with that of conjunctive questions: $?P\vee?Q$ has to be ruled out on semantic or pragmatic grounds (Hoeks and Roelofsen 2019) and we most naturally analyse alternative questions as $P \vee Q$ and open disjunctive questions as $?(P \vee Q)$ (Roelofsen and Farkas 2015). Either way, we do not straightforwardly get transparency.

In fact, by switching to inquisitive semantics we might even have lost our explanation of presupposition projection in assertions, at least within the theories of presupposition projection that we focus on here. When we derive why the Transparency Theory predicts a presupposition to the effect that $\neg p$ is true to be satisfied in the second disjunct of $p \vee q$, the following property is crucial: $p \vee \top = p \vee \neg p$, where $\top$ is a tautology (cf. Schlenker 2008). This property does not hold in inquisitive semantics (here Ω is the set of all worlds and $\mathcal{P}$ maps sets to their powerset).

(45) a. $P \vee \top = \{s \mid s \vdash p\} \cup \mathcal{P}(\Omega) = \mathcal{P}(\Omega)$
b. $P \vee \neg P = \{s \mid s \vdash p\} \cup \{s \mid s \vdash \neg p\} = ?P \neq \mathcal{P}(\Omega)$

A similar problem will arise when we derive local contexts in the manner of Schlenker (2009) or when we derive trivalent connectives in the manner of George (2014). The natural way around this issue is to look at the logical properties of $!(P \vee Q)$ rather than $P \vee Q$; in inquisitive semantics, $!(P \vee Q)$ is essentially a classical disjunction. However, since the presence of the operator ! is assumed to be specific to assertions, predicating our derivation of projection facts upon it makes it so that our analysis does not extend to questions. Equivalently, we could say that what matters to presupposition projection is not full equivalence between propositions, but rather *informational* equivalence (the notion obtained by ignoring inquisitiveness); however, this would lead to undesirable predictions and require further amendments to the system (e.g., all polar questions would become equivalent, so that we might end up predicting an absence of projection from polar nuclei).

Categorial theories The final approach to question semantics that I will mention is the categorial approach. In categorial theories, questions are assigned a complex functional type that depends on the kind of question we are looking at. Thus a polar question might have a type isomorphic to a proposition, or a more complex type (e.g. $((st)st)st$ for Krifka (2001)), while a *who* question would be a one-place property of individuals (type $e \to s \to t$ or $s \to e \to t$). The appeal is, among other things, that this allows for finer-grained distinctions than other theories. In particular, categorial theories always avoid yes/no-symmetry.

The main challenge for categorial theories is that questions behave in a way that suggests they all have the same type: they can be relatively freely coordinated, and a number of attitude verbs can embed any kind of question. In order to solve the former problem, proponents of categorial approaches usually resort to the reductionist view in its higher-order type variant (see in particular Krifka 2001 and Xiang 2021), with the problems that we have seen. Alternatively, they fall back to one of the above theories to account for coordination and embedding. Categorial theories thus do not help us towards understanding our phenomenon.

6.3 A proposal: questions as trivalent inquisitive predicates

6.3.1 Introduction and background

In this section, I propose a novel view of what kind of object questions denote, which is rich enough to let us derive presupposition projection patterns while avoiding the issues we have encountered in our earlier attempts.

While this choice is not a necessity, it will make the motivation for the proposal and the exposition clearer if we adopt a trivalent view of propositions and presuppositions; in Section 6.4.2 we will see how the system can function with bivalent propositions as well. The trivalent approach consists in reifying presuppositions by assuming that there are three truth values: 0, to which propositions map worlds where they are false; 1, to which propositions map worlds where they are true; and #, to which propositions map worlds where their presuppositions are not satisfied. Presupposition projection can then be derived from a trivalent semantics for logical operators (Peters 1979; Beaver and Krahmer 2001; George 2014).

It is customary in the formal semantics literature to entertain two perspectives on propositions at the same time: propositions are functions from worlds to truth values but they are also sets of worlds. In a bivalent setting, these two domains are isomorphic and the mapping between them is trivial. In a trivalent setting, the domain of functions from worlds to (trivalent) truth values is richer than the domain of sets of worlds, so in what follows I will distinguish between the functional type $s \to t$ of trivalent propositions and the set type $\{s\}$

of sets of worlds, as that will make certain things clearer. The operator $\mathbb{1}$ maps a trivalent proposition to its truth set:

(46) $\mathbb{1}(p) := \{w \mid p(w) = 1\}$

Finally, the trivalent perspective lets us define an operator that maps a trivalent proposition to "its" presupposition. This operator will be denoted as π; it is defined in (47a). Note that what π returns is a set of worlds (there is no sense in which the presupposition of a proposition could itself have a presupposition). The identity in (47b) will be useful to have in mind.[29]

(47) a. $\pi(p) := \{w \mid p(w) \in \{0, 1\}\}$
b. $\forall p.\, \pi(p) = \mathbb{1}(p) \cup \mathbb{1}(\neg p)$

6.3.2 Trivalent homogeneous polar questions

The system I present here is a trivalent theory of polar questions. We consider a polar question $?p$ to be a predicate over sets of worlds (construed as epistemic states). This predicate is potentially undefined: we take $?p$ to be defined at a state s if and only if s settles p, i.e. s supports p or $\neg p$. Thus:

(48) $$?p = \lambda s. \begin{cases} 1 & \text{if } s \vdash p, \\ 0 & \text{if } s \vdash \neg p, \\ \# & \text{in all other cases.} \end{cases}$$

Here p is a trivalent proposition. As before, $s \vdash p$ is read as "s supports p" and means that p maps all worlds in s to 1. Thus $s \vdash \neg p$ will hold when s maps all worlds in s to 0. If s does not support the presuppositions of p — that is, if there are #-worlds in s — neither $s \vdash p$ nor $s \vdash \neg p$ can hold. There are therefore two reasons why we can have $?p(s) = \#$: (i) if s includes both 1-worlds and 0-worlds, i.e. s does not settle p; (ii) if s includes #-worlds, i.e. s does not satisfy the presuppositions of p. These two reasons are of course compatible with each other.

It is helpful to compare our denotation to what is assumed in inquisitive semantics. Inquisitive semantics also assumes that questions are predicates over sets of worlds, but it does not distinguish positive and negative answers. Thus, the standard inquisitive treatment of polar questions, when converted from set notation to functional notation, is essentially what is given in (49):

(49) $$?_{\text{inq}}p = \lambda s. \begin{cases} 1 & \text{if } s \vdash p \text{ or } s \vdash \neg p, \\ 0 & \text{otherwise.} \end{cases}$$

The system I propose here can therefore be seen as an extension of the inquisitive account of polar questions. Unlike the basic inquisitive account, it assigns to them a yes/no-asymmetric denotation: $?p$ and $?(\neg p)$ are different objects.[30] Also unlike the basic inquisitive account, it is integrated within a trivalent logic, making it easy to develop a theory of

[29] Here $\neg$ denotes trivalent negation: it maps 0 to 1, 1 to 0, and # to #.

[30] The basic idea of adding yes/no-asymmetry to inquisitive semantics has already been explored by Roelofsen and Farkas (2015), but their system is conceptually quite different. Roelofsen and Farkas (2015) propose a two-dimensional theory where $?p$ has its usual inquisitive denotation on the ordinary dimension (essentially $\{p, \neg p\}$). On the second dimension, called the "highlighting" dimension, $?p$ essentially denotes p.

presupposition projection based on it, as we will see. One may wonder how the usage of the third value (#) in question denotations relates to its usual usage: is the system based on the idea that questions presuppose something? Tentatively, I submit that the fact that undecided states are mapped to # can be related to the pragmatic effect of questions, which require the addressees to address them before the conversation can move on. Thus questions arguably do presuppose that certain states will not be undecided: specifically, they presuppose that *future Common Grounds* will have addressed them (see also footnote 38). At this point, however, I do not have a definitive answer as to how to interpret the distinction between positive answers; I certainly do not intend here to develop a system where questions can be "true" or "false". While I believe that this distinction could be linked to certain pragmatic phenomena, this will not be discussed here. Instead, we can presume for now that the distinction is relevant to nothing other than question coordination. The most conservative interpretation would be that it is merely a formal device.[31]

The inquisitive denotation can be defined in terms of the denotation under our proposal; we call it the "domain" of a question. Intuitively, as in inquisitive semantics, the domain is the set of states where the question is resolved. Also as in inquisitive semantics (cf. (43)), we can define the "alternatives" of a question, the least specific states where the question is resolved, which correspond intuitively to the possible answers. For a polar question $?p$, $\textsc{alt}(?p)$ will be $\{\mathbb{1}(p), \mathbb{1}(\neg p)\}$, i.e. the denotation that is generally assumed in Hamblin-Karttunen semantics.[32]

(50) $\textsc{dom}(Q) := \{s \mid Q(s) \in \{0,1\}\}$

(51) $\textsc{alt}(Q) := \{s \mid s \in \textsc{dom}(Q) \wedge \neg\exists s' \in \textsc{dom}(Q).\, s \subsetneq s'\}$

(52) a. $\textsc{dom}(?p) = \{s \mid s \vdash p\} \cup \{s \mid s \vdash \neg p\}$
b. $\textsc{alt}(?p) = \{\mathbb{1}(p), \mathbb{1}(\neg p)\}$

Finally, we define an operator that maps a question to its *informational commitment*, the union of all states where the question is defined. Intuitively, this is (a superset of) what the questioner takes for granted (the intuition is that the questioner is assuming that the question can be resolved; we will formalize it in section 6.3.4).

(53) $\textsc{info}(Q) := \bigcup \textsc{dom}(Q)$

In the case of a polar question $?p$, $\textsc{info}$ will return the union of p's truth worlds and p's falsity worlds, leaving out the #-worlds. This is essentially the presupposition of p:

(54) $\textsc{info}(?p) = \pi(p)$

This identity will let us derive the fact that a polar question has the same presuppositions

[31] Under the proposal presented here, answers to a complex question will all be assigned a certain category (positive or negative). For instance, $p \wedge q$ will be a positive answer to $?p \wedge ?q$, while $p \wedge \neg q$ will be a negative one. As an anonymous reviewer points out, there is no clear intuition on how to make this distinction beyond simple polar questions that we could evaluate the proposal against. My claim is that this particular way of dividing up "positive" and "negative" answers will lead to an adequate account of question coordination and presupposition projection, which is enough to justify it. Unless we ascertain that another phenomenon involves a more fine-grained distinction or a different one, there is no reason to be concerned about the premise of categorizing answers beyond what is intuitively obvious.

[32] In the possible extension to constituent questions based on $\exists_G$ that we are going to discuss in Section 6.4.4, $\textsc{alt}$ will specifically return (roughly) the Hamblin-Karttunen answers to the question, as opposed to propositions corresponding to the partition cells in the sense of Groenendijk and Stokhof (1984). For polar questions, there is no difference between Hamblin answers and partition cells.

as the corresponding assertion.

To give a concrete illustration of how all the definitions I just introduced work, consider (55a). This sentence presupposes that John once smoked. We therefore assign to it, under a trivalent view of presuppositions, the denotation in (55b). The corresponding polar question in (56a) is assigned the denotation in (56b). In (57) we see that the domain of the question is the union of the set of 1-states and the set of 0-states; these are all the subsets of the set of 1-worlds ($\mathbb{1}(p)$) and the set of 0-worlds ($\mathbb{1}(\neg p)$) respectively. The domain therefore has two maximal elements, the set of 1-worlds and the set of 0-worlds, and these constitute the alternatives as seen in (58). Finally, the informational commitment, seen in (59), is the union of the domain, which is equivalent in our case to the union of the alternatives: the set of worlds where John used to smoke and either stopped or still does, i.e. the set of all the worlds where John used to smoke.

(55) a. p: John stopped smoking.

b. $$p = \lambda w. \begin{cases} 1 & \text{if John used to smoke and stopped in } w, \\ 0 & \text{if John used to smoke and still does in } w, \\ \# & \text{if John never smoked in } w. \end{cases}$$

(56) a. $?p$: Did John stop smoking?

b. $$?p = \lambda s. \begin{cases} 1 & \text{if John used to smoke and stopped throughout } s, \\ 0 & \text{if John used to smoke and still does throughout } s, \\ \# & \text{in all other cases.} \end{cases}$$

(57) $$\begin{aligned} \textsc{dom}(?p) &= \{s \mid ?p(s) = 1\} \cup \{s \mid ?p(s) = 0\} \\ &= \mathcal{P}(\{w \mid p(w) = 1\}) \cup \mathcal{P}(\{w \mid p(w) = 0\}) \\ &= \mathcal{P}(\mathbb{1}(p)) \cup \mathcal{P}(\mathbb{1}(\neg p)) \\ &= \mathcal{P}(\{w \mid \text{John used to smoke and stopped in } w\}) \\ &\quad \cup \mathcal{P}(\{w \mid \text{John used to smoke and still does in } w\}) \end{aligned}$$

(58) $$\begin{aligned} \textsc{alt}(?p) &= \{\mathbb{1}(p), \mathbb{1}(\neg p)\} \\ &= \left\{ \begin{matrix} \{w \mid \text{John used to smoke and stopped in } w\}, \\ \{w \mid \text{John used to smoke and still does in } w\} \end{matrix} \right\} \end{aligned}$$

(59) $$\begin{aligned} \textsc{info}(?p) &= \bigcup \textsc{dom}(?p) \\ &= \left(\bigcup \mathcal{P}(\mathbb{1}(p))\right) \cup \left(\bigcup \mathcal{P}(\mathbb{1}(\neg p))\right) \\ &= \mathbb{1}(p) \cup \mathbb{1}(\neg p) \\ &= \pi(p) \\ &= \{w \mid \text{John used to smoke in } w\} \end{aligned}$$

6.3.3 Defining the connectives

The extension of the system to conjunctions and disjunctions of polar questions is straightforward. The type we assign to questions is conjoinable, and we are going to assume that *and* and *or* apply to it directly. In a trivalent theory, there are several sensible candidates for conjunction and disjunction. Peters (1979) argues that presupposition projection patterns suggest we should adopt left/right-asymmetric connectives. These connectives are known as *Peters connectives* or *Middle Kleene* connectives (in contrast to Strong and Weak Kleene;

cf. Beaver and Krahmer 2001); their definitions are given in the tables below (rows represent the left-hand argument).[33]

(60)

$\wedge$	0	1	#
0	0	0	0
1	0	1	#
#	#	#	#

$\vee$	0	1	#
0	0	1	#
1	1	1	1
#	#	#	#

Our assumption will be that question coordination obeys Middle Kleene rules. From this we immediately get a denotation for conjunctive and disjunctive questions, and we can compute their domain and therefore their resolution conditions (under the assumption that a question is resolved when a proposition in its domain is known to be true).

Here we see that the distinction between positive and negative answers, which has played no role so far, will crucially affect the denotation of coordinated questions.

6.3.4 Deriving presupposition projection

To give an account of presupposition projection in our system, we need a version of Stalnaker's bridge for questions, or in other words, we need to specify what in a question's denotation makes it felicitous to ask or not in a given context.

As mentioned above, we are going to identify the presuppositions of a question with the output of INFO. The motivation for this view goes as follows: in general, questions do not modify the set of worlds in the context set, but they restrict what continuations are allowed, in the sense that unless the participants contribute an assertion that addresses the question, the conversation will be at an impasse.[34] It seems natural then to consider that a speaker should not ask a question if it is not established (in the Common Ground) that the participants are capable in principle of contributing such an assertion. It should be necessary, then, that when a question is asked, the Common Ground entail that there exists a proposition that is true at the actual world and addresses the question. Going back to our system, the speaker should only utter Q if it is established that there exist states in DOM(Q) that contain the actual world, or equivalently that the Common Ground supports INFO(Q).

We have already seen that INFO($?p$) $= \pi(p)$. Thus, we predict a polar question to presuppose the same thing as the corresponding assertion. This is a well-known fact, and is sometimes used as a test or even a definition for presuppositions.[35]

6.3.5 Conjunctive questions

We can now see what our system predicts conjunctive questions to mean.

Denotation and presuppositions The denotation of a conjunction of questions $?p \wedge ?q$ is as follows:

[33] George (2014) motivates the MK connectives by showing how they can be derived from classical bivalent logic together with some general assumptions about definedness conditions.

[34] This is of course an idealization where we ignore, among other things, the possibility of avowing complete or partial ignorance ("I don't know", "Probably", etc.), the possibility of rejecting the question ("Who cares?"), the possibility of providing an implicit answer ("A: Did John come? B: He was sick."), the existence of rhetorical questions, and so on.

[35] I am ignoring here the various pragmatic presuppositions that are associated to asking a question: that the facts have not been established before in the discourse, that the answer is accessible, etc.

$$?p\wedge?q = \lambda s.\begin{cases} 1 & \text{if } s \vdash p \text{ and } s \vdash q, \\ 0 & \text{if } s \vdash \neg p \text{ or if } s \vdash p \text{ and } s \vdash \neg q, \\ \# & \text{in all other cases.} \end{cases} \tag{61}$$

From this we can compute the domain and the alternatives. The domain being the downwards closure of the alternatives, it is sufficient to give the alternatives:

$$\text{ALT}(?p\wedge?q) = \{\mathbb{1}(\neg p), \mathbb{1}(p \wedge \neg q), \mathbb{1}(p \wedge q)\} \tag{62}$$

The alternatives are mutually exclusive. If p has no presupposition, they form a three-cell partition of the logical space. The three cells are essentially those I called the tripartition when discussing H/K semantics, and we have seen that they correspond intuitively to the possible answers to our presuppositional conjunctive question (4).

We can compute the predicted presupposition of a conjunctive question, as follows:

$$\begin{aligned} \text{INFO}(?p\wedge?q) &= \bigcup \text{DOM}(?p\wedge?q) \\ &= \bigcup \text{ALT}(?p\wedge?q) \\ &= \mathbb{1}(\neg p) \cup (\mathbb{1}(p) \cap \mathbb{1}(q)) \cup (\mathbb{1}(p) \cap \mathbb{1}(\neg q)) \\ &= \mathbb{1}(\neg p) \cup (\mathbb{1}(p) \cap (\mathbb{1}(q) \cup \mathbb{1}(\neg q))) \\ &= \mathbb{1}(\neg p) \cup (\mathbb{1}(p) \cap \pi(q)) \\ &= (\mathbb{1}(\neg p) \cup \mathbb{1}(p)) \cap (\mathbb{1}(\neg p) \cup \pi(q)) \\ &= \pi(p) \cap (\mathbb{1}(\neg p) \cup \pi(q)) \end{aligned} \tag{63}$$

The end result is an intersection of two terms. The first term is $\pi(p)$, the presupposition of p: for $?p\wedge?q$ to be felicitous, all worlds in the context therefore have to satisfy the presuppositions of p. The second term means that worlds in the context set should either make p false or satisfy the presuppositions of q. Essentially, this is equivalent to saying that they should satisfy the material conditional $p \to \pi(q)$. These two conditions on the context are exactly the ones that have been argued since Karttunen (1973; 1974) to describe presupposition projection in conjunctive assertions, and that I have argued in Section 6.1 to also describe presupposition projection in conjunctive questions. This is not surprising, because once we have derived the tripartition, this projection pattern follows in a fairly theory-agnostic way.

The main difference with H/K semantics and inquisitive semantics here is that the way we derive the tripartition is significantly different. In these other theories, a conjunction of polar questions $?p\wedge?q$ most naturally corresponds to what I called the quadripartition, and the tripartition can only be derived by assuming a left/right-asymmetric structure ($?(p\wedge?q)$ in inquisitive semantics), if it can be derived at all. Within the present proposal, we get the tripartition from a left/right-symmetric structure because our conjunction is left/right-asymmetric in its semantics.

Contrast with earlier accounts The reason that earlier accounts derive the quadripartition is of course not just that it naturally falls out from the formalism, but also that it is thought to be an empirically adequate characterisation of what conjunctive questions mean in general, absent presuppositional effects. It is of course desirable that we maintain this empirical coverage.[36]

There are several ways in which we can derive the quadripartition in our system. First, we can assume that *and* may also denote a Weak Kleene trivalent conjunction. Second, we can assume there is an operator T, defined as in (64), that makes all the answers positive. If so, then $\text{T}(?p)\wedge ?q$ (or $\text{T}(?p) \wedge \text{T}(?q)$) would have the alternatives of the quadripartition. Finally, we can assume that whenever we think we observe the quadripartition, what we are actually seeing is two separate questions, the coordination of which takes place at the level of attitude or speech acts. The result of high-level coordination is indeed virtually indistinguishable in most cases from the quadripartition: if I know the answer to both conjuncts, I know the answer to the quadripartition. This is the basis of what we have called reductionist theories (e.g. Hirsch 2017; Xiang 2021).

(64) $$\text{T}(Q) := \lambda s. \begin{cases} 1 & \text{if } Q(s) = 1 \text{ or } Q(s) = 0 \\ \# & \text{if } Q(s) = \# \end{cases}$$

Whatever solution we adopt, we end up with a system where conjunctive questions can have the resolution conditions of either the tripartition or the quadripartition. We assume that examples of presuppositional filtering involve the tripartition; recall that because it is yes/no-symmetric, the quadripartition cannot be the basis of an analysis of this phenomenon. When there is no presuppositional trigger that forces the tripartition, conjunctive questions should in principle be ambiguous. This raises the following question: why do we generally observe quadripartitive readings, and are there counterexamples?

I would like to claim that, in an appropriate context, conjunctive questions can have the answerhood conditions of the tripartition even when no presupposition is present to force such a reading. An example is given in (65). (65a) has the form $?p\wedge ?q$, and a fully knowledgeable speaker may answer it felicitously with $\neg p$, even though that should be under-informative relative to the quadripartition. This tends to show that the tripartition does exist. The tripartition also predicts that $\neg q$ as an answer should be judged uncooperative. What might in fact be the case, even though the judgment is not clear, is that $\neg q$ as an answer implicates p. While the present account does not offer an immediate explanation for this fact, it would be impossible to derive it from the quadripartition due to its yes/no-symmetry; the tripartition does not have this problem.[37] Additionally, the quadripartition would predict that if p is replaced by p', which is equivalent to $\neg p$, as in (65b), then we should see no difference. What we observe is that p' becomes an uncooperative answer.

(65) Context: *Mary has been working in an underground garage. John has been outside, so he knows the weather, and he was the last person to use the grill, so he knows about the charcoal.*

a. Mary: Is the weather nice outside, and is there any charcoal left?

[36] H/K semantics and inquisitive semantics generate a variety of other readings, as we have seen. In particular, there is a polar reading ($\{p \wedge q, \neg(p \wedge q)\}$) and a reverse tripartition ($\{\neg q, p \wedge q, \neg p \wedge q\}$). Without further assumptions, our system generates none of these either. In principle, the polar reading can be generated if we assume that *and* can take scope below ?, or if we have an operation that collapses the 0-alternatives into one. The reverse tripartition can be generated if we assume that *and* can also denote a right-to-left rather than left-to-right Middle Kleene connective, or equivalently that it can combine either with the first or the second conjunct first. The reverse tripartition is unattested and, pending further investigation, my judgment is that the polar reading is not available either (cf. footnote 24), so these extra assumptions are not necessary.

[37] One could in particular use the fact that $p \wedge \neg q$ is a good answer and entails $\neg q$. Asserting $\neg q$ always raises the contextual probability of $p \wedge \neg q$ relative to the other two cells, from which an inference that $p \wedge \neg q$ holds might follow as a relevance implicature.

(i) John: It's raining outside. ($\neg p$)
$\not\rightsquigarrow$ There is no charcoal left.
$\not\rightsquigarrow$ There is some charcoal left.

(ii) ? John: There is no charcoal left. ($\neg q$)
$\overset{?}{\rightsquigarrow}$ The weather is nice. (p)

b. # Mary: Is it raining outside, and is there any charcoal left?
John: It's raining outside. ($p' \equiv \neg p$)

A related example is given in (66a). Here the context guarantees that the quadripartition cannot possibly be addressed by the participants in the conversation, while the tripartition can. A conjunctive question is perfectly felicitous, suggesting that the issue it raises is the tripartition. An advocate of the quadripartition will of course argue that we have not independently established that it is infelicitous to ask a conjunctive question that cannot be fully addressed, but if conjunctive questions always denote the quadripartition and not being able to tell between some of the cells is fine, there is no reason why reversing the order of the conjuncts as in (66b), or replacing the left-hand nucleus with its negation as in (66c), leads to infelicity.

(66) Context: *Mary has applied for a grant. The decision is due to come by physical mail; none of the participants has any other way of finding out about it. Someone asks:*

a. Has the mail arrived, and did Mary get her grant?
b. #Did Mary get her grant, and has the mail arrived?
c. #Has the mail failed to arrive, and did Mary get her grant?

Thus, the tripartition can be observed in non-presuppositional cases. What both our examples have in common is that the reason the speaker is not interested in the quadripartitive issue is clear from the context: in (65), it is more information than they need to establish whether they can carry out their plan or not, and in (66a), it is more information than the listener can possibly be knowledgeable about.[38] This suggests that the reason that most examples found in the literature lend themselves to quadripartitive readings is purely pragmatic: in the absence of an explicit context, the quadripartitive reading is more natural because it requires less specific assumptions on the speaker's motivations. Indeed, if p and q are arbitrary, unrelated propositions, it is hard to see why a speaker would be interested in the tripartition formed from $?p\wedge?q$, while being interested in the quadripartition is just being interested in both $?p$, and $?q$, possibly for different reasons.

It should be noted as a counterpoint that most speakers seem to judge matrix conjunctive questions somewhat degraded out-of-the-blue. This is puzzling if conjunctive questions can easily denote the quadripartition, which I have just argued should be easy to accept without context. It would be consistent, however, with the view that conjunctive questions always denote the tripartition, and that the quadripartition is an artefact of high-scope coordination, which might be unavailable or marked in matrix questions.

[38] One might in fact argue that both (65) and (66a) do involve presupposition filtering, but with pragmatic presuppositions. In (65), the presupposition that the second question is relevant to some greater issue is not satisfied in the context: it only matters whether there is charcoal if the weather is nice, as the goal is to take out the grill. In (66a), similarly, the presupposition that knowledge of the answer to the second question is accessible to the participants is not satisified, but what is true is that if the mail has arrived, participants have a chance knowing whether Mary got her grant. This is consistent with the interpretation of the trivalent system offered in Section 6.3.2: a polar question $?q$ on its own presupposes that future Common Grounds will entail q or $\neg q$, but when it is in the second conjunct of $?p\wedge?q$, this presupposition can be filtered, and it is fine if the question is uninteresting or unknowable when p is false.

Another counterpoint is that apparent quadripartitive readings of conjunctive questions do not necessarily constitute strong evidence that these questions denote the quadripartition. First, it is in general felicitous for answerers to provide more information than the questioner was explicitly asking for. (67) is a felicitous dialogue and is not taken to provide evidence that B's utterance is a semantic answer to A's question. Thus the fact that a conjunctive question $?p\wedge?q$ can generally be answered with $\neg p \wedge \neg q$ is not necessarily incompatible with a tripartitive analysis. In contrast, being less precise than was asked is not felicitous in general, which is why (65) is puzzling from a quadripartitive perspective.

(67) A: Is the weather nice outside?
B: Yes, but there is no charcoal left.

6.3.6 Disjunctive questions

Disjunctive questions work very similarly.

Denotation and presuppositions We derive for them the following denotation and alternatives:

(68) $$?p\vee?q = \lambda s. \begin{cases} 1 & \text{if } s \vdash p \text{ or if } s \vdash \neg p \text{ and } s \vdash q, \\ 0 & \text{if } s \vdash \neg p \text{ and } s \vdash \neg q, \\ \# & \text{in all other cases.} \end{cases}$$

(69) $$\textsc{alt}(?p\vee?q) = \{\mathbb{1}(p), \mathbb{1}(\neg p \wedge q), \mathbb{1}(\neg p \wedge \neg q)\}$$

The predicted presupposition is derived as follows:

(70) $$\begin{aligned} \textsc{info}(?p\vee?q) &= \bigcup \textsc{dom}(?p\vee?q) \\ &= \bigcup \textsc{alt}(?p\vee?q) \\ &= \mathbb{1}(p) \cup (\mathbb{1}(\neg p) \cap \mathbb{1}(q)) \cup (\mathbb{1}(\neg p) \cap \mathbb{1}(\neg q)) \\ &= \mathbb{1}(p) \cup (\mathbb{1}(\neg p) \cap (\mathbb{1}(q) \cup \mathbb{1}(\neg q))) \\ &= \mathbb{1}(p) \cup (\mathbb{1}(\neg p) \cap \pi(q)) \\ &= (\mathbb{1}(p) \cup \mathbb{1}(\neg p)) \cap (\mathbb{1}(p) \cup \pi(q)) \\ &= \pi(p) \cap (\mathbb{1}(p) \cup \pi(q)) \end{aligned}$$

The only difference with conjunction is that we have p instead of $\neg p$ in the second term. Thus, we again predict the presuppositions of p to project, but the material conditional that should hold throughout the context set is now $\neg p \rightarrow \pi(q)$ rather than $p \rightarrow \pi(q)$. Once again, these contextual requirements are exactly what Karttunen (1973; 1974) argues should be derived for disjunctive assertions, and what I argued in Section 6.1 to be the observed pattern in disjunctive questions.

Contrast with earlier accounts: the open reading Per (69), we predict that the issue raised by a disjunctive question involves three mutually incompatible possibilities, corresponding essentially to the answer set $\{p, \neg p \wedge q, \neg p \wedge \neg q\}$. This is intuitively a good match for the open reading of disjunctive questions, which also has three good answers. However, our partition cells do not exactly align with what Roelofsen and Farkas (2015) derive in their inquisitive account (and what we derived earlier in our H/K reconstruction of it). Roelofsen and Farkas (2015) assume the structure $?(p \vee q)$ for open disjunctive questions, from which we get the equivalent of $\{p, q, \neg p \wedge \neg q\}$. Thus our alternatives exhibit a left/right asymmetry that is not found in their account. Concretely, they predict that the question is resolved when the participants know the positive answer to either disjunct, even though they might be ignorant about the other one, while in our system the participants need to know for sure what the truth value of p is.

It is not clear which prediction is more correct: an example like (71) shows some left/right asymmetry in the resolution conditions, as we predict: if the answerer says q, we draw the inference that $\neg p$, which would be consistent with the idea that the partition cell being identified here really is $\neg p \wedge q$. With a symmetric approach, it is formally impossible to account for the contrast. The characterization that the question has not been resolved if one is ignorant about John might nevertheless be too strict.

(71) A: Did John↗ arrive or did Mary?↗
  a. B: John arrived. ⇝ Mary might or might not have arrived.
  b. B: Mary arrived. $\overset{?}{⇝}$ (B thinks) John didn't arrive.

We can add that, in some other cases, no asymmetry can clearly be detected:[39]

(72) A: Can I call you tomorrow at 3pm↗ or on Friday at 1pm?↗
  a. B: You can call me tomorrow. ⇝ B does not want A to call them on Friday.
  b. B: You can call me on Friday. ⇝ B does not want A to call them tomorrow.

This could suggest that, in at least some cases, $\{p, q, \neg p \wedge \neg q\}$ really is the answer set that we want. We can generate it within our system by assuming that disjunction can denote a Strong Kleene rather than Middle Kleene connective, or that it always does. If that is the case, we expect to observe presupposition filtering not just from the right-hand disjunct through the left-hand one, but also the other way round. Recall that the data is unclear on this point: we do observe a left-right contrast, but it is much less clear than for conjunction, and in the case of assertions, Rothschild (2011) among others has argued that the data does not support an asymmetric view of filtering. In a way, our discussion whether the resolution conditions of open disjunctive questions are left/right-asymmetric or not reflects the earlier debate whether presupposition filtering in disjunctive assertions is left/right-asymmetric or not. We will leave both questions open, but venture that they should receive a common answer.

Contrast with earlier accounts: the closed reading Disjunctive questions also have a closed reading (or alternative question reading). In an alternative question like (73), the two propositions p and q inside each disjunct are presupposed to be exhaustive ($p \vee q$ holds) and exclusive ($p \wedge q$ does not hold). In H/K semantics, alternative questions are generally assigned the denotation $\{p, q\}$; the equivalent move in inquisitive semantics is to assign to them the structure $p \vee q$. Alternative questions can be resolved by asserting either of the two disjuncts, p or q.

(73) Did John↗ arrive or did Mary?↘

To derive a two-element answer set from the denotation of $?p \vee ?q$, it seems natural to

[39] I thank an anonymous reviewer for this example.

assume a mechanism through which a presupposition that a question can be answered positively can be generated, for instance an operator POS:

(74) $\text{POS}(Q) := \lambda s : Q(s) = 1.\, Q(s)$

(75) $\text{ALT}(\text{POS}(?p \vee ?q)) = \{p, \neg p \wedge q\}$.

If that is our account of alternative questions, we correctly predict that they have two possible answers, and that p and q are good answers in a context where exhaustiveness ($p \vee q$) and exclusivity ($\neg(p \wedge q)$) hold (since in such a context, q and $\neg p \wedge q$ are equivalent). We also derive the exhaustiveness presupposition. We do not derive the exclusivity presupposition in any way, however, so we have to assume that it arises from a separate mechanism, perhaps as an effect of focus structure in the question nucleus: cf. Section 6.4.5 for discussion. Under the Strong Kleene variant of the theory, the alternatives are $\{p, q\}$, as in the traditional account, but again the exclusivity presupposition is not derived.

Finally, recall that in inquisitive semantics and H/K semantics, a structure like $?p \vee ?q$ will result in the answer set $\{p, \neg p, q, \neg q\}$, which has to be ruled out one way or another, as there is no reading of disjunctive questions that looks like it. The system here does not generate this problematic answer set.

6.3.7 Interim conclusion

At this point, I have given an account of conjunctive questions and disjunctive questions such that the presupposition projection patterns that we observed in Section 6.1 can be easily derived. I have argued that the predicted resolution conditions for conjunctive and disjunctive questions are attested, and that those predicted by earlier accounts can also be generated, with only a few, simple extra assumptions.

The analysis offers a fully parallel account of conjunctions and disjunctions of polar questions. They are both assumed to have the structure that the surface syntax suggests, that is to say, $?p \square ?q$, where $\square$ is the (generalized) truth-conditional binary connective that you would expect, in its Middle Kleene variant.

Furthermore, assuming that propositions receive a classical trivalent denotation (type $s \to t$, where t is the type of trivalent truth values), we can give a single entry for "and" and for "or" that applies to both questions and propositions. We just have to make them polymorphic truth-conditional connectives, in a way similar to what Partee and Rooth (1983) propose (with the added twist that we understand truth values to be trivalent):

(76) a. $[\![\text{and}]\!] = \lambda Q_{\alpha \to t}.\, \lambda P_{\alpha \to t}.\, \lambda x_\alpha.\, P(x) \wedge Q(x)$
b. $[\![\text{or}]\!] = \lambda Q_{\alpha \to t}.\, \lambda P_{\alpha \to t}.\, \lambda x_\alpha.\, P(x) \vee Q(x)$
(type $\forall \alpha.\, (\alpha \to t) \to (\alpha \to t) \to \alpha \to t$ in both cases)

Thus all our desiderata as listed in Section 6.2.3 are met.

Of course, the theory presented here is probably not the only way to achieve these desiderata, and slight reformulations as well as radically different (possibly dynamic) theories could probably be offered. I do believe that they would involve some form of the following idea: there is some notion of a positive and a negative answer to a question, even beyond polar questions (cf. footnote 31), and the connectives should interact with this distinction in a similar manner to the way they interact with propositions. At an even more abstract level, the way we analyse a question should be more parallel to the way we analyse related propositions than it currently is. As I discussed in Section 6.2 and demonstrated for the case of answer set semantics, the lack of such properties in earlier accounts makes it very difficult to deal with the data presented here.

6.4 Further predictions and challenges for the trivalent inquisitive theory of questions

This final section explores in more detail the theory we have just developed: we first revisit some data points that we have set aside so far, and show how we are able to deal with them, and then briefly show we could develop the theory into a general theory of questions.

6.4.1 The effect of *or not*

Recall that, even though an alternative question using *or not* such as (77) has the same resolution conditions as the corresponding polar question, when such questions occur in a coordinated structure, they do not license presupposition filtering (cf. (9b)/(18b)).

(77) Did John come or not?

Within our framework, *or not* questions are naturally analysed as instantiating the structure $?p{\vee}?(\neg p)$, which we will abbreviate as $\tilde{?}p$. Thus, they receive the denotation in (78). This makes *or not* questions completely similar to polar questions in terms of their domain, their alternatives, and their informational commitment, as seen in (79).

(78) $$\tilde{?}p = \lambda s. \begin{cases} 1 & \text{if } s \vdash p \text{ or } s \vdash \neg p, \\ \# & \text{otherwise.} \end{cases}$$

(79) a. $\textsc{dom}(\tilde{?}p) = \{s \mid s \vdash p \vee s \vdash \neg p\}$
b. $\textsc{alt}(\tilde{?}p) = \{\mathbb{1}(p), \mathbb{1}(\neg p)\}$
c. $\textsc{info}(\tilde{?}p) = \pi(p)$

The main way in which *or not* questions are predicted to differ from polar questions is that they behave differently with respect to coordination. Consider for instance $\tilde{?}p{\wedge}?q$:

(80) $$\tilde{?}p{\wedge}?q = \lambda s. \begin{cases} 1 & \text{if } s \vdash p \wedge q \text{ or } s \vdash \neg p \wedge q, \\ 0 & \text{if } s \vdash p \wedge \neg q \text{ or } s \vdash \neg p \wedge \neg q, \\ \# & \text{in all other cases.} \end{cases}$$

(81) a. $\textsc{alt}(\tilde{?}p{\wedge}?q) = \{p \wedge q, \neg p \wedge q, p \wedge \neg q, \neg p \wedge \neg q\}$
b. $\textsc{info}(\tilde{?}p{\wedge}?q) = \pi(p) \cap \pi(q)$

Unlike $?p{\wedge}?q$, $\tilde{?}p{\wedge}?q$ evokes a four-cell partition. Furthermore, the presuppositions of the second conjunct are predicted to project no matter what. Thus, we predict (82) (repeated from (9b)) to be infelicitous.

(82) #Is Syldavia a monarchy or not, and is the Syldavian monarch a progressive?

6.4.2 Local contexts, triviality, and bivalent propositions

In Section 6.1, we stated a generalization about the observed data in terms of local contexts. The point was to also capture certain redundancy effects, seen in (83a) and (83b) (repeated from (26a) and (26b)), and in (84a) and in (84b) (repeated from (27a) and (27b)).

(83) a. #Is Ann in Paris and is she in France?
b. #Is Ann in Paris and is she in London?

(84) a. #Is Ann away from Paris or is she in France?
b. #Is Ann away from Paris or is she in London?

The infelicity of all these examples is predicted by a generalization in terms of local contexts (the local context of q in $?p \wedge ?q$ entails p, the local context of q in $?p \vee ?q$ entails $\neg p$), together with a triviality constraint: no constituent should be trivially true or false in its local context.

Here, we have been adopting a trivalent theory of presuppositions, and we have not mentioned local contexts so far. It is nevertheless straightforward to derive local contexts that are exactly the ones we need, by applying Schlenker's (2009) procedure to our question denotations.[40] The facts above will then be correctly predicted.

If we derive local contexts in this way, since they can also be used to predict presupposition projection, as shown by Schlenker (2009), we no longer need trivalent connectives at the propositional level. Thus the Middle Kleene connectives would exclusively serve for questions, and the present proposal would come down to the fact that question coordination obeys some form of Middle Kleene logic, without tying it as closely to presupposition projection. Since the empirical predictions are the same in the cases that we have been looking at so far, the choice between these two approaches has to be made on other grounds.

It should be noted that at least as far as these particular facts are concerned, we do not absolutely need a theory of local contexts. The various infelicitous examples listed above run afoul of plausible pragmatic constraints on question denotation:

- (83a) would be banned by a global redundancy constraint: we predict it to be equivalent to its first conjunct on its own.
- (83b) is predicted to only have 0-alternatives, as its 1-alternative is a contradiction. It is plausible that this is pragmatically deviant: the yes/no-asymmetries in polar questions identified in the literature (Büring and Gunlogson 2000; Roelofsen and Farkas 2015, a.o.) suggest that the positive answer has to have some special significance in the discourse context.[41]
- Within existing H/K approaches, (84a) would be banned by a unique-answer constraint: the two disjuncts overlap in natural contexts. In our system, the disjunction is equivalent to "Is Ann away from Paris or not?", and could be banned by a global redundancy constraint similarly to (83a).
- A unique-answer constraint would also ban (84b) in H/K theory. In our system, the disjunction is in fact equivalent to its first disjunct, "Is Ann away from Paris?", so a global redundancy constraint would ban it.

[40] Schlenker (2009) applies his procedure to bivalent propositions, but the procedure is mostly agnostic with respect to the type of what it is looking at, and it is straightforward to adapt it to functions from states to trivalent values. The one thing we need to specify is how contextual restriction works with states: we take it that if the context set is C, the set of contextually permissible states is the set of subsets of C.

[41] I think the distinction between 1-alternatives and 0-alternatives would be helpful in a theory of question bias, but this is not the place to develop this idea.

6.4.3 Embedding and the relation between questions and propositions

Predicates like *know* can embed both questions, as in (85a), and propositions, as in (85b). It is of course desirable to analyse these two examples with a single entry for *know*, while accounting for the major difference in meaning between them.

(85) a. John knows whether Mary came.
b. John knows that Mary came.

Uegaki (2019) compares several approaches, including:

- the *Q-to-P approach*: assuming that *know* embeds propositions, and that there is an operator that converts questions into propositions;
- the *P-to-Q approach*: assuming that *know* embeds questions, and that there is an operator that converts propositions into questions;
- the *uniformitarian approach*: giving the same kind of denotation to propositions and questions.

He argues that the Q-to-P approach cannot account for certain phenomena that are easy to analyze under the P-to-Q and uniformitarian approaches.[42]

At any rate, our theory can in principle be made compatible with all three strategies. Here I will briefly describe an implementation of P-to-Q.[43] The P-to-Q approach can be implemented in our system in a fairly conservative way, mostly following the same lines as what would be done in answer set semantics. We can assume, as we have been doing so far, that propositions have type $s \to t$, where t is the type of trivalent truth values. The following conversion operator converts a proposition to a question (type $\{s\} \to t$, where $\{s\}$ is the type of (non-trivalent) sets of worlds):

(86) $$\textsc{that}(p) := \lambda s. \begin{cases} 1 & \text{if } s \vdash p, \\ \# & \text{otherwise.} \end{cases}$$

Note that $\textsc{that}(p) = \textsc{pos}(?p)$. Then, the following entry for *know* will deliver the appropriate truth conditions and presuppositions for (85a) and (85b).

(87) $$[\![\text{know}]\!] = \lambda w. \lambda Q. \lambda x : w \in \textsc{info}(Q).\ \text{Dox}^{w}(x) \vdash \textsc{ans}_S(Q)(w)$$

where $\textsc{ans}_S$ is an answerhood operator that computes the answer to a question:

(88) $$\textsc{ans}_S(Q)(w) = \left(\bigwedge\{p \mid p \in \textsc{alt}(Q), p(w) = 1\}\right) \wedge \left(\bigwedge\{\neg p \mid p \in \textsc{alt}(Q), p(w) = 0\}\right)$$

[42] For discussion of the differences between P-to-Q and uniformitarian approaches, see Roelofsen 2019 and references therein.

[43] Q-to-P can be implemented through answerhood operators; cf. Section (97). The uniformitarian approach can be implemented by assigning to propositions the same type as questions. The simplest approach consists in assigning to propositions the same denotation as to the corresponding polar question, and distinguishing them only at the highest level (above any connective), by assuming either that polar questions are embedded under POS (then the entry for *know* would presuppose that we are in a 1-state), or that assertions are embedded under something similar to THAT (then the entry for *know* needs not be changed).

Applied to a polar question $?p$, $\textsc{ans}_S$ returns p if p is true and $\neg p$ if p is false. The definition does not have to be as complicated as it is to get this result, but it will also extend to constituent questions, as we will see below. In the propositional case, we correctly capture that "know that p" presupposes that p is true, as $\textsc{info}(\textsc{that}(p)) = \mathbb{1}(p)$.

6.4.4 Extending the proposal to constituent questions

I will only sketch an extension of the system to constituent questions here, in order to show what it might look like. I will be making a specific assumption on how constituent questions should be formalized, though alternative approaches can certainly be pursued. There are of course a number of issues that might prove challenging to tackle, and I will mention a few.

Trivalent existential quantification To extend our system to constituent questions like (89), we need to give them a denotation that is a trivalent predicate upon sets of worlds, as for polar questions.

(89) Who came?

The assumption that we make here, then, is that constituent questions are existentially quantified-into polar questions. This is consistent with the fact that *wh*-words are crosslinguistically similar to indefinites, and it is going to deliver the alternatives that we want. We can implement the idea as follows:

(90) $[\![\text{who}]\!] = \lambda P.\, \lambda s.\, \exists x \in \textbf{human}.\, ?(P(x))(s)$

(91) $[\![(89)]\!] = \lambda s.\, \exists x \in \textbf{human}.\, ?(\textbf{came}(x))(s)$

For the sake of perspicuity, we will ignore the restriction to humans and adopt the following notation:

(92) $?x.\, P(x) \equiv \lambda s.\, \exists x.\, ?P(x)(s)$

What the denotation we derive actually comes down to depends on what kind of trivalent existential quantifier we adopt. I propose to adopt the lazy trivalent existential quantifier defined by Beaver and Krahmer (2001) and George (2014), which is described in (93).[44]

(93) $$\exists x \in D.\, P(x) = \begin{cases} 0 & \text{if for all } y \in D, P(y) = 0, \\ 1 & \text{if there is } y \in D \text{ such that } P(y) = 1, \\ \# & \text{in all other cases.} \end{cases}$$

We then derive the following domain, alternatives, and informational commitment:

(94) $\textsc{dom}(?x.\, P(x)) = \{s \mid [\forall x.\, s \vdash \neg P(x)] \vee [\exists x.\, s \vdash P(x)]\}$

(95) $\textsc{alt}(?x.\, P(x)) = \{\bigcap_x \mathbb{1}(\neg P(x))\} \cup \{\mathbb{1}(P(x)) \mid x\}$

(96) $\textsc{info}(?x.\, P(x)) = [\exists x.\, P(x)] \vee [\forall x. \neg P(x)]$

[44] A natural alternative would be a universally-projecting existential quantifier, such that $\exists x.\, \phi_x$ is defined only if ϕ_x is defined for all x. This would make the alternatives to the question congruent with its strongly exhaustive answers, which is unwelcome because there is no obvious way to derive weakly exhaustive and mention-some answers from strongly exhaustive answers.

The alternatives are the *mention-some answers* to (89), plus the negative answer. The presupposition we predict for a constituent question involving a presuppositional predicate, such as (97), is a roughly existential presupposition similar to what George (2014) predicts for existential quantification: at least one person has a dog and is walking it, or everyone has a dog and nobody is walking their dog.[45]

(97) Who's walking their dog?

Answerhood operators for constituent questions Much of the discussion in the literature about the semantics of constituent questions is applicable to our system as long as we are able to redefine the various notions of answers that have been established in the literature. Since the alternatives that we derive are very similar to what is commonly assumed in Hamblin-Karttunen semantics, this is relatively straightforward to accomplish.

As we have seen, our alternatives already correspond to the mention-some answers to the question. A second kind of answer is found in the strongly exhaustive answers, which provide full information about the truth of the alternatives. The strongly exhaustive answer to "Who came?", for instance, is the proposition asserting that the people who came came, and that the people who did not come did not come. We have already seen above, in (88), the definition of an operator computing the strongly exhaustive answer, and an entry for *know* based on it. Finally, some question-embedding predicates like *surprise* are thought to be sensitive to weakly exhaustive answers. The weakly exhaustive answer to "Who came?" is the proposition that the people who came came, with no information about the fact that nobody else did. Thus, (98) means that John was surprised that the people who came came, as opposed to being surprised that the people who did not come did not come.

(98) It surprised John who came.

It is also straightforward to define an operator that computes the weakly exhaustive answer through a conjunction, as seen in (99). Here we are exploiting the difference between alternatives that are mapped to 0 and to 1; this enables us to go around the fact that our alternatives include an extra element ("nobody came") relative to what is generally assumed in H/K semantics, where it is often assumed that the negative answer is not a semantic answer (e.g. Dayal 1996).

(99) $\text{ANS}_W(Q)(w) := \bigwedge\{p \mid p \in \textsc{alt}(Q) \wedge Q(p) = 1 \wedge p(w) = 1\}$

To conclude, the system proposed here can be extended to constituent questions in a way that is fairly close to H/K semantics and lets us replicate its basic ingredients, with the main difference (which we can get around if we wish) lying in our assumption that the negative answer to a constituent question is an actual semantic alternative. The parallelism with H/K semantics makes it so that much of the discussion found in the literature is applicable to the trivalent system.

[45]The matter of how presuppositions project from questions beyond simple polar cases has received relatively little attention in the literature and is subject to debate. Schwarz and Simonenko (2018) argue that while universal projection is generally observed in constituent questions (as noted in earlier work, e.g., Abruśan 2011), it can be obviated in certain contexts. In a paper published after the present one was written, Theiler (2021) proposes a generalization whereby what we observe is universal projection over an answer set that can be smaller than the H/K question denotation under certain conditions. Integrating Theiler's observations and proposal with the present framework is left to future work.

6.4.5 Some potential challenges

A problem with plurals and Dayal's presupposition A common alternative to the definition of weakly exhaustive answers found above involves maximization rather than conjunction, following Dayal (1996). If the domain includes Ann and Bill, and *who* is taken to also range over plural individuals, under standard assumptions a constituent question like (89) should denote the set in (100) in H/K semantics.

(100) {Ann came, Bill came, Ann and Bill came}

Because the set is closed under conjunction, the weakly exhaustive answer is always in the set, and it can be recovered without a conjunction operation. Instead, it can be defined as the maximal element (under entailment) within the true answers:

(101) $\textsc{ans}_W(Q_{HK})(w) := \iota p.\, p \in Q \wedge p(w) \wedge [\forall p' \in Q.\, p'(w) \rightarrow p \vDash p']$

The expression above is only defined if there is a single p verifying the scope of ιp. It is frequently assumed that questions in fact presuppose that the actual world makes the expression above defined: this is Dayal's presupposition, also known as the strongest-answer presupposition. Under this assumption, among other niceties, the difference in presuppositions between "Which student came?" (assumed to be as in (102)) and "Which students came?" (assumed to be as in (100)) follows naturally.

(102) {Ann came, Bill came}

Unfortunately, this line of thought is not directly transferable to our system. Indeed, because we define alternatives as maximal elements under set inclusion, we are committed to the fact that alternatives are logically independent. Thus, we cannot have an alternative set that looks like (100). The set of positive alternatives for (89) can only look like (102), whether the *wh*-word ranges over singular individuals only or over both singular and plural individuals.

This has at least two consequences. First, we need to define weakly exhaustive answers in the conjunctive manner above. Second, the uniqueness presupposition of "Which student came?" cannot be derived from a more general presupposition of questions. Instead, it has to be analyzed as a presupposition specific to *which* questions.[46]

Deriving the exclusivity inference in alternative questions A related issue is that, as we have seen above, we do not derive the exclusivity presupposition of alternative questions. Recall that an alternative question or closed disjunctive question carries the presupposition that exactly one of the disjuncts is true. The presupposition can be further decomposed into exhaustiveness (at least one disjunct is true: $p \vee q$) and exclusivity (the two disjuncts are not both true: $\neg(p \wedge q)$). With the denotation that is traditionally assumed in H/K semantics, $\{p, q\}$, both exhaustiveness and exclusivity can be derived from a strongest-answer constraint in the spirit of Dayal (1996). In our system, assuming that a closed disjunctive question denotes $\textsc{pos}(?p \vee ?q)$ (as predicted by a Middle Kleene view of disjunction), the

[46] Champollion, Ciardelli, and Roelofsen (2017) defend a localist account of *which*'s presupposition within the framework of inquisitive semantics, where the same problem occurs. Decoupling the presuppositions of *which* questions from answerhood operators is also argued for by Uegaki (2020) and Hirsch and Schwarz (2019) within H/K semantics. An alternative solution to the issue consists in making singular and plural *which* questions different again through extensions of the inquisitive system; see for instance the dynamic inquisitive proposal of Dotlačil and Roelofsen (2020).

alternatives are $\{p, \neg p \wedge q\}$. A strongest- or unique-answer constraint, applied to our alternatives, derives exhaustiveness but not exclusivity. Worse than that, it is in fact impossible to construct a constraint that derives exclusivity from $\{p, \neg p \wedge q\}$, because the desired presupposition ($\neg(p \wedge q)$) is not expressible as a function of p and $\neg p \wedge q$. The Strong Kleene view of disjunction is more promising: the alternatives for POS($?p \vee ?q$) are $\{p, q\}$, and a unique-answer constraint will derive the exclusivity presupposition. Adding a unique-answer presupposition will not affect our analysis of conjunctive cases (where the alternatives form a partition anyway). We will need to assume that it is not operative in open disjunctive questions, which do not seem to presuppose anything, but this problem is already present in earlier accounts. If we want instead to maintain the Middle Kleene view, which, as we have seen, helps explain certain left/right asymmetries, one solution might be to adopt a similar approach as in the case of *which* questions, consisting in the proposal that the individual alternatives entail the exclusivity inference. However, unlike in the case of *which*, there is no independent motivation for the move.

6.5 Conclusion

In this long paper, we have identified a striking parallelism between the presupposition projection patterns observed when a trigger is in the second of two coordinated polar questions, and when a trigger is in the second of two coordinated assertions. We have seen that existing views on the semantics of coordinated questions struggle to explain the projection pattern in itself, and that they would have an even harder time explaining the parallelism, because they do not analyse coordinated questions in a way that adequately relates them to coordinated assertions. The system proposed here, a trivalent extension of inquisitive semantics, solves the initial problem of deriving the projection facts by making polar questions more closely related to assertions. Through this proposal, we also make polar questions yes/no-asymmetric, and we give a uniform account of question conjunction and question disjunction — two issues that have proven problematic for some theories of questions. While we do not have the space to explore further implications here, a natural next step will be to identify which other phenomena the resulting system might help us with, such as, possibly, question bias. Additionally, while I think I have shown here a defect in existing theories of questions taken as-is, there certainly exist other ways of extending extant theories, following the same basic conceptions that I followed when extending inquisitive semantics, which would deliver the same welcome results; we will have to see whether some of them might not improve on the present system in some respects.

6.A Appendix: Formal companion to section 6.2.4

This appendix contains a formal development that demonstrates the problems encountered when one combines answer set semantics with either the Transparency Theory or trivalent theories, as cursorily explained in Section 6.2.4. We will derive the fact that the observed presupposition filtering pattern is expected in conjunctive questions if they are analysed as the tripartition, but not if they are analysed as the quadripartition. Meanwhile, the observed pattern in disjunctive questions is not derived in any natural way.

Note that throughout the Appendix we work with question denotations, rather than with declarative sentences containing embedded questions, and therefore the results might only directly apply to matrix questions. In principle, at least under the Transparency Theory, we could predict different filtering patterns under certain embeddings (in particular, we might expect more filtering predictions). While I do not think this would lead to different results in practice, no attempt is made to prove it here.

When the proof of a result is omitted, it is because it is immediate.

6.A.1 Equivalence relations on questions

The Transparency Theory requires a notion of equivalence; we define several natural ones below. As in the main text, we identify questions with Hamblin denotations (sets of propositions). We will use Q, Q', etc. to name our abstract variables representing questions. In order to show that the results extend to Karttunen's approach, we will also deal with Karttunen denotations (functions from worlds to sets of propositions), which we will refer to as $\hat{Q}$, $\hat{Q}'$, etc. The following relations let us map each kind of denotation into the other:

Definition 1 (Relation between Hamblin and Karttunen denotations)**.** For a question Q, we define:

$$\hat{Q} := \lambda w. \{p \in Q \mid p(w)\}$$

As long as Q does not contain a logical contradiction, the following relation lets us reverse the mapping:

$$Q = \bigcup_w \hat{Q}(w)$$

Note that it is not true in general that all Karttunen questions can be derived from a Hamblin question (i.e., that they can be written as $\hat{Q}$ for some Q). However, all the questions that we consider here are representable as Hamblin questions.

In the main text, we left the notion of equivalence we were using somewhat implicit. Here we define several reasonable notions of equivalence. These notions are all relativised to a context set C (they are notions of contextual equivalence). The purpose of this wealth of definitions is to show that the issues outlined in the main text do not depend on specific implementational choices within the general framework of H/K semantics. In particular, whether we adopt H-equivalence, which is the most natural approach from the perspective of a Hamblin-style theory, or K-equivalence, which is the most natural approach from the perspective of a Karttunen-style theory, does not matter, showing that our discussion applies to Hamblin- and Karttunen-style accounts equally.

For the notions based on a presuppositional answerhood operator, I implement a variety of ways of dealing with the unique-answer presupposition. The first approach treats it essentially as an entailment, a second one only considers *worlds* where it is true, and a third one only considers *contexts* where it is true.[47]

Non-contextual versions of the non-presuppositional notions can be obtained by taking C to be the set of all possible worlds.

Definition 2 (Pointwise contextual restriction)**.**

(i) For a proposition p and a context set C, $p_{|C}$ is the function from C to truth values that is identical to p at all points (the *contextual restriction* of p to C).

(ii) For a question Q and a context set C, the contextual restriction of Q to C is $Q_{|C}$, where:

$$Q_{|C} := \{p_{|C} \mid p \in Q\}$$

[47]Both the second and third approach are reminiscent of the definition of Strawson entailment proposed by von Fintel (1999).

Definition 3 (Hamblin equivalence). Q and Q' are *H-equivalent* relative to context set C iff $Q_{|C} = Q'_{|C}$.

Definition 4 (Karttunen equivalence). Q and Q' are *K-equivalent* relative to context set C iff:

$$\forall w \in C.\, [\hat{Q}(w)]_{|C} = [\hat{Q}'(w)]_{|C}$$

Definition 5 (Unique-answer answerhood operator). We define:

$$\textsc{ans}_U := \lambda Q.\, \lambda w.\, \iota p.\, p \in Q \wedge p(w)$$

Definition 6 (Unique-answer equivalence: entailed presupposition). Q and Q' are *U-equivalent* relative to a context set C iff:

(i) the unique-answer presuppositions of Q and Q' are equivalent in C,

(ii) calling C' the subset of C where the presuppositions are met, we have:

$$\forall w, w' \in C'.\, \textsc{ans}_U(Q)(w)(w') = \textsc{ans}_U(Q')(w)(w')$$

Definition 7 (Unique-answer Strawson equivalence — world-level version). Q and Q' are *W-equivalent* relative to a context set C iff for any world w in C such that the unique-answer presuppositions of both Q and Q' are true at w, we have:

$$\forall w' \in C.\, \textsc{ans}_U(Q)(w)(w') = \textsc{ans}_U(Q')(w)(w')$$

Definition 8 (Unique-answer Strawson equivalence — context-level version). Q and Q' are *C-equivalent* relative to a context set C iff for any subset C' of C where the unique-answer presuppositions of both Q and Q' are true, we have

$$\forall w, w' \in C'.\, \textsc{ans}_U(Q)(w)(w') = \textsc{ans}_U(Q')(w)(w')$$

Result 1 (H-equivalence and K-equivalence are almost the same thing).

(i) If C is a context set and Q and Q' are two questions, then if Q and Q' are H-equivalent in C, they are K-equivalent in C.

(ii) If C is a context set and Q and Q' are two questions that do not contain any proposition that is false throughout C, then if Q and Q' are K-equivalent in C, they are H-equivalent in C.

Result 2 (Order of the answer-based equivalences by strength). If C is a context set and Q and Q' are two questions, then if Q and Q' are U-equivalent in C, they are W-equivalent in C, and if they are W-equivalent in C, they are C-equivalent in C.

Result 3 (K-equivalence is stronger than unique-answer equivalences). If C is a context set and Q and Q' are two questions, then if Q and Q' are K-equivalent in C, they are U-equivalent and therefore also W-equivalent and C-equivalent in C.

Result 4 (Collapse of the equivalences). If C is a context set and Q and Q' are two questions whose unique-answer presuppositions are satisfied in C, then for any pair of letters α, β within H, K, U, W, and C, Q and Q' are α-equivalent in C iff they are β-equivalent in C.

6.A.2 The Transparency Theory

Here we derive the announced results under Schlenker's (2008) Transparency Theory. Schlenker (2008) proposes that in a context C, a presuppositional clause is acceptable if and only if the proposition being presupposed is *transparent* in the clause's position. Transparency is defined as follows (some details are simplified or left implicit; the reader is referred to Schlenker 2008):

Definition 9 (Transparency; Schlenker 2008)**.** Let $\alpha\beta\gamma$ be a sentence where β is an embedded clause, and δ be another clause denoting a proposition d. d is *transparent* in the position of β if and only if for any completion γ' that makes $\alpha\beta\gamma'$ well-formed, and for any clause β', the two sentences $\alpha\beta'\gamma'$ and $\alpha(\delta \text{ and } \beta')\gamma$ are contextually equivalent in C.

The definition of Transparency depends on a notion of contextual equivalence; it is straightforward to apply it to questions as long as we have defined equivalence over them. Thus we can in principle define H-transparency, K-transparency, and so on in terms of the definitions above.

The relations between our notions of equivalence immediately translate into relations between our notions of transparency:

Result 5 (Relations between the transparencies)**.**

(i) H-transparency implies K-transparency.

(ii) K-transparency implies U-transparency.

(iii) U-transparency implies W-transparency.

(iv) W-transparency implies C-transparency.

We can now derive the results that we are interested in. The first one below establishes that under the quadripartitive account, no presupposition filtering should ever be observed when the trigger is in the second conjunct of a conjunctive question.

Result 6 (The quadripartition)**.** Let C be a context set, and p a proposition such that there exists a proposition q_0 that is not related by contextual entailment to p or $\neg p$.[48] A proposition d is H-transparent (as well as K/U/W/C-transparent) in the position of q in $\text{C}[?p] \wedge \text{C}[?q]$ if and only if C supports d.

Proof. The direction "If C supports d, then d is transparent" is immediate.

For the other direction, it suffices to show that if d is C-transparent, then C supports d, as C-transparency is the weakest form of transparency. Note that the definedness condition of ANS_U is always met if the answers form a logical partition as they do here, and therefore we do not need to consider subcontexts at all.

Assume then that d is C-transparent. We will write $Q(q) = \text{C}[?p] \wedge \text{C}[?q]$; the fact that d is C-transparent means that $\text{ANS}_U(Q(q))(w)$ is equivalent in C to $\text{ANS}_U(Q(d \wedge q))(w)$ for any q and any $w \in C$ (taking β' to be q, and γ' to be the empty string).

It follows from the existence of q_0 that p is not trivially true or false in C. Let then w, w' be two worlds such that $p(w) = 1$ and $p(w') = 0$. We have $\text{ANS}_U(Q(\top))(w) = p \wedge \top = p$ and $\text{ANS}_U(Q(\top))(w') = \neg p \wedge \top = \neg p$.

[48] This condition should clearly be met in any realistic example. Possibly a weaker condition would suffice to validate the result.

- If $d(w) = d(w') = 0$, then $\text{ANS}_U(Q(d))(w) = p \wedge \neg d$, from which it follows that $p \wedge \neg d$ is equivalent to p, i.e. that p entails $\neg d$. We also have $\text{ANS}_U(Q(d))(w') = \neg p \wedge \neg d$, from which it follows similarly that $\neg p$ entails $\neg d$. The only way these two entailments can hold is if d is a trivial falsehood. But then, we have $\text{ANS}_U(Q(d \wedge q_0))(w) = p \wedge \neg(d \wedge q_0)$, which is equivalent to p, so p is equivalent to $\text{ANS}_U(Q(q_0))(w)$, which is either $p \wedge q_0$ or $p \wedge \neg q_0$. This contradicts the assumption that p and $q_0/\neg q_0$ are not related by entailment.
- If $d(w) = 1$ and $d(w') = 0$, then $\text{ANS}_U(Q(d))(w) = p \wedge d$ and $\text{ANS}_U(Q(d))(w') = \neg p \wedge \neg d$. It follows that p entails d and that $\neg p$ entails $\neg d$. Hence p is equivalent to d. Then, we have $\text{ANS}_U(Q(d \wedge q_0))(w') = \neg p \wedge \neg(d \wedge q_0)$, which is equivalent to $\neg p$, so $\neg p$ is equivalent to $\text{ANS}_U(Q(q_0))(w')$, which is either $\neg p \wedge q_0$ or $\neg p \wedge \neg q_0$. It follows that $\neg q_0$ or q_0 entails $\neg p$, against assumption.
- If $d(w) = 0$ and $d(w') = 1$, the same reasoning as in the previous case applies, replacing any occcurrence of p with $\neg p$ and vice versa, and replacing w' with w.

The only remaining possibility is that $d(w) = d(w') = 1$. Since w can be any p-world and w' can be any $\neg p$-world, d has to be a trivial truth in C. □

The next result shows that if conjunctive questions denote the tripartition, then we predict the observed pattern of presupposition filtering in the second conjunct.

Result 7 (The tripartition). Let C be a context set, and p a proposition such that there exists a proposition q_0 that is not related by contextual entailment to p or $\neg p$. A proposition d is H-transparent (as well as K/U/W/C-transparent) in the position of q in $\text{C}[?p \wedge \text{C}[?q]]$ if and only if C supports the material conditional $p \to d$.

Proof. We define: $Q(q) := \text{C}[?p \wedge \text{C}[?q]] = \{\neg p, p \wedge \neg q, p \wedge q\}$.

Assume that C supports $p \to d$, and let q be an arbitrary proposition. $p \wedge d \wedge q$ is equivalent in C to $p \wedge q$, and this relation can equivalently be written as $[p \wedge d \wedge q]_{|C} = [p \wedge q]_{|C}$. Moreover, we have $p \wedge \neg(d \wedge q) = (p \wedge \neg d) \vee (p \wedge \neg q)$, which is equivalent to $p \wedge \neg q$ (as $p \wedge \neg d$ is a contextual contradiction). It follows that $Q(q)$ and $Q(d \wedge q)$ are H-equivalent in C. Then, for any completion γ, $Q(q)\gamma$ and $Q(d \wedge q)\gamma$ are also H-equivalent in C.[49] Therefore, d is H-transparent in the position of q in $Q(q)$, and it is also K/U/W/C-transparent.

As for the other direction, assume that d is C-transparent in the position of q in $Q(q)$ in C. As before, it follows from the existence of q_0 that p is not trivially true or false in C. Let us then take $w \in C$ such that $p(w) = 1$. If $d(w) = 0$, since $\text{ANS}_U(Q(\top))(w) = p$ and $\text{ANS}_U(Q(d))(w) = p \wedge \neg d$, p and $p \wedge \neg d$ are contextually equivalent, which is equivalent to saying that C supports $p \to \neg d$. Assume without loss of generality (as we could replace q_0 by $\neg q_0$) that $q_0(w) = 1$. Then, we have $\text{ANS}_U(Q(d \wedge q_0))(w) = p \wedge (\neg d \vee q_0)$. Since p entails $\neg d$, it also entails $\neg d \vee q_0$, and therefore $\text{ANS}_D(Q(d \wedge q_0))(w)$ is equivalent to p. By C-transparency, p is thus equivalent to $\text{ANS}_U(Q(q_0))(w)$, i.e. p is equivalent to $p \wedge q_0$, or equivalently, p contextually entails q_0, against assumptions. Therefore this case is impossible, and $d(w) = 1$. Since this holds for any w such that $p(w) = 1$, C supports $p \to d$. □

Moving on to disjunction, the next result shows that no filtering is predicted to be possible with H/K/U-transparency. The subsequent result is an extension to W-transparency (that an additional condition is needed is essentially a bug in the definition). A final result shows that C-transparency derives a degenerate pattern where everything is transparent.

[49] We would need to define an explicit fragment to properly prove this fact, but I think it is reasonable to take it for granted.

Result 8 (No filtering in alternative questions (H/K/U))**.** Let C be a context set. A proposition d is H/K/U-transparent in the position of q in $?p \vee ?q$ in C if and only if C supports d.

Proof. The direction "If C supports d, then d is transparent" is immediate.

We define $Q(q) :=?p \vee ?q = \{p, q\}$. Assume that d is U-transparent in C. The unique-answer presupposition of $Q(\top)$ is equivalent to $\neg p$, from which it follows that the unique-answer presupposition of $Q(d)$ is equivalent to $\neg p$, i.e. that $p \veebar d$ is equivalent to $\neg p$ ($\veebar$ represents an exclusive disjunction). This can be verified to be equivalent to the fact that d is trivially true in C. □

Result 9 (No filtering in alternative questions (W))**.** Let C be a context set, and let d be a proposition that does not contextually entail p.[50] d is W-transparent in the position of q in $?p \vee ?q$ if and only if C supports d.

Proof. Once again, define $Q(q) :=?p \vee ?q = \{p, q\}$.

The direction "If C supports d, then d is transparent" is immediate.

Assume that d is W-transparent. Let w be a world such that $p(w) = 0$ and $d(w) = 1$ (such a world exists by assumption). $\textsc{ans}_U(Q(\top))(w) = \top$ and $\textsc{ans}_U(Q(d)) = d$, so d is contextually equivalent to $\top$. □

Result 10 (C-equivalence derives a degenerate pattern in alternative questions)**.** Let C be a context set, and let d be a proposition. d is C-transparent in the position of q in $?p \vee ?q$.

Proof. Once again, define $Q(q) :=?p \vee ?q = \{p, q\}$.

Take q to be an arbitrary proposition, and call C' the set of worlds in C such that $p \vee (d \wedge q)$ and $\neg(p \wedge q)$ are true. C' is the biggest subset of C such that the unique-answer presupposition of both $Q(q)$ and $Q(d \wedge q)$ is met. What we need to prove is that $\textsc{ans}_U(Q(q))$ and $\textsc{ans}_U(Q(d \wedge q))$ define the same two-place predicate over C', or equivalently that for any $w \in C'$, both operators return equivalent propositions when applied to w, with equivalence being taken in C'.[51] Take then w in C'.

- If $p(w) = 1$, then $\textsc{ans}_U(Q(q))(w) = \textsc{ans}_U(Q(d \wedge q))(w) = p$.
- If $p(w) = 0$, from the definition of C' we have $d(w) = q(w) = 1$. It follows that $\textsc{ans}_U(Q(q))(w) = q$ and $\textsc{ans}_U(Q(d \wedge q))(w) = d \wedge q$. Due to the way C' is defined, both q and $d \wedge q$ are equivalent to $\neg p$ in C' and therefore to one another, as desired.

Thus, d is C-transparent. □

These results are essentially preserved if we look at $\{p, q, \neg(p \vee q)\}$ rather than $\{p, q\}$; the proofs are omitted to save space.

[50] If d contextually entails p, then the unique-answer presupposition of $\{p, d \wedge q\}$ will be contradictory, whatever q is. d is actually W-transparent in this case, but not in an interesting way.

[51] The fact that we look at equivalence in C' rather than in C here constitutes the crucial difference between C-equivalence and W-equivalence.

6.A.3 Trivalent theories: a compositional approach

We now turn to trivalent theories. The simplest way of adapting trivalent theories to our problem is to consider that conjunction and disjunction of propositions are trivalent, and that question conjunction and disjunction just do what they usually do in H/K semantics, except that questions are sets of *trivalent* propositions. Below we define such a system, called the Simple Trivalent Answer set Theory (STAT). It is straightforward to verify that under STAT, essentially the same results as with the Transparency Theory obtain: the conjunctive case can be dealt with through the tripartition, while the disjunctive case remains puzzling.[52]

Concretely, assume the following:

Definition 11 (STAT).

(i) Questions are sets of trivalent propositions.

(ii) Question conjunction is pointwise Middle Kleene conjunction.

(iii) Question disjunction is set union.

(iv) Questions are felicitous only if all their answers return 0 or 1 at all worlds in the context.

Result 11 (Predictions of STAT). Under the assumptions given in Definition 11:

(i) If a conjunctive question (schema: $?p \wedge ?q$) denotes the quadripartition $\{p \wedge q, p \wedge \neg q, \neg p \wedge q, \neg p \wedge \neg q\}$, it should presuppose $\pi(p)$ and $\pi(q)$.

(ii) If a conjunctive question (schema: $?p \wedge ?q$) denotes the tripartition $\{p \wedge q, p \wedge \neg q, \neg p\}$, it should presuppose $\pi(p)$ and $p \rightarrow \pi(q)$.

(iii) Whether a disjunctive question (schema: $?p \vee ?q$) denotes $\{p, q\}$ or $\{p, q, \neg p \wedge \neg q\}$, it should presuppose $\pi(p)$ and $\pi(q)$.

In order to deal with disjunction, one might think that case (iv) in Definition 11 should be relaxed to an existential presupposition: *at least one* answer should be defined at a given world in the context. We can in fact unify the constraint with the independent assumption, generally made in H/K theories, that questions presuppose that one of their answers is true:

Definition 12 (Existential variant of STAT).

(iv') Questions presuppose that at least one answer is defined and true at each world in the context.

The observed case of filtering is immediately predicted, but we derive no order effects.

[52] An entirely parallel result obtains in the case of the Simple Dynamic Answer set Theory (SDAT):

Definition 10 (SDAT).

(i) Questions are sets of dynamic propositions denoting Context Change Potentials (CCPs), as defined in Heim 1983.

(ii) Question conjunction is pointwise CCP composition.

(iii) Question disjunction is set union.

(iv) Questions are felicitous only if all their answers are defined at the context.

Li (2019) proposes a system of this kind.

6.A.4 Trivalent deployment of the connectives

Another way of extending trivalent theories to questions is to apply the methodology of "Peters-Kleene deployment", as described by George (2014), to derive trivalent meanings for the question connectives from their bivalent meanings. Below we provide an implementation of George's idea, and derive the same results as with the Transparency Theory: presupposition filtering is predicted for the tripartition but not for the quadripartition as far as conjunction is concerned, and not at all as far as disjunction is concerned (only the case of $\{p, q\}$ is discussed, as using $\{p, q, \neg p \wedge \neg q\}$ instead does not make a difference).

The idea behind Peters-Kleene deployment is as follows: assume that you want to derive how potential undefinedness in p will project in the environment $F(p)$. What you know is the specification of F in the bivalent world (i.e., you know what $F(p)$ is for a total proposition p). Now define the *repair set* of a trivalent proposition p: it is the set of all propositions that agree with p wherever p is defined.

Definition 13 (Repair set of a trivalent proposition)**.** If p is a trivalent proposition, the *repair set* of p is written as p^R and given by:

$$p^R := \{p' \in \{0,1\}^\Omega \mid \forall w.\, p(w) \in \{0,1\} \to p'(w) = p(w)\}$$

(Here Ω is the set of possible worlds, and therefore $\{0,1\}^\Omega$ is the set of total/bivalent propositions.)

From the repair set of p, we derive the deployment of F, F^D. Unlike $F(p)$, $F^D(p)$ is potentially defined for some trivalent inputs p. Those inputs are those where, no matter how p is repaired to form a bivalent proposition p', $F(p')$ is the same.

Definition 14 (Deployment of a functor)**.** If F is a function whose input is a bivalent proposition, the *deployment* of F is written as F^D and given by:

(i) If there is an output X such that for all $p' \in p^R$, $F(p') = X$, then $F^D(p) = X$.

(ii) Otherwise, $F^D(p) = \#$.

In our case, F will be the function from q to the denotation of the question schematized as $?p\wedge?q$ or $?p\vee?q$. The desired result is that even if q is potentially undefined, $F(q)$ might be defined in some cases.

Applying this methodology is most naturally done within a Karttunen-style account. Thus, the output of K will be a Karttunen question (type $s(st)t$). The definition of deployment above presupposes a notion of equality on the outputs of F; the natural choice in our case is the usual definition of set equality.[53]

We can now prove the results promised above:

Result 12 (Deployment of conjunction (quadripartition))**.** For total propositions p and q, define $F_p(q)$ to represent the Karttunen denotation of a quadripartitive conjunctive question:

$$F_p(q) := \lambda w.\lambda r_{st}.\, r \in \{p \wedge q, p \wedge \neg q, \neg p \wedge q, \neg p \wedge \neg q\} \wedge r(w)$$

Let p be a total proposition that is true in at least two worlds and false in at least two worlds.[54] $F^D_p(q)$ is defined if and only if q is total.

[53]Note that we could relativize everything to a context set, as is done in the definition of K-equivalence, while preserving all the results.

[54]As before, this condition should be met in any realistic example.

Proof. It is clear that if q is total, $F_p^D(q)$ is defined (it is just $F_p(q)$).

Assume that q is not total, i.e. that there is w such that $q(w) = \#$. Take $q' \in q^R$ such that $q'(w) = 1$, and $q'' \in q^R$ that is exactly the same as q', except that $q''(w) = 0$.

- If $p(w) = 1$, then $F_p(q')(w) = \{p \wedge q'\}$ and $F_p(g'')(w) = \{p \wedge \neg q''\}$. Let w' be another world where p is true: we have $[p \wedge q'](w') = q'(w')$ and $[p \wedge q''](w') = q''(w')$. By construction, $q'(w') \neq q''(w')$, so $F_p(q')(w) \neq F_p(g'')(w)$.
- If $p(w) = 0$, then $F_p(q')(w) = \{\neg p \wedge q'\}$ and $F_p(q'')(w) = \{\neg p \wedge \neg q''\}$. Let w' be another world where p is false: we have $[\neg p \wedge q'](w') = q'(w')$ and $[\neg p \wedge q''](w') = q''(w')$. By construction, $q'(w') \neq q''(w')$, so $F_p(q')(w) \neq F_p(g'')(w)$.

Either way, $F_p^D(q)$ is not defined. By contraposition, if $F_p^D(q)$ is defined, q is total. □

Result 13 (Deployment of conjunction (tripartition))**.** For total propositions p and q, define $F_p(q)$ to represent the Karttunen denotation of a tripartitive conjunctive question:

$$F_p(q) := \lambda w. \lambda r_{st}. r \in \{p \wedge q, p \wedge \neg q, \neg p\} \wedge r(w)$$

Let p be a total proposition that is true in at least two worlds. $F_p^D(q)$ is defined if and only if all #-worlds for q are 0-worlds for p.[55]

Proof. Assume that q is only undefined at worlds where p is false. Take q', q'' in q^R, and let w be a world.

- If $p(w) = 0$, then $[p \wedge q'](w) = [p \wedge q''](w) = 0$, and $[p \wedge \neg q'](w) = [p \wedge \neg q''](w) = 0$
- If $p(w) = 1$, then q is defined at w and $q'(w) = q''(w) = q(w)$. If $q(w) = 1$, we have $[p \wedge q'](w) = [p \wedge q''](w) = 1$ and $[p \wedge \neg q'](w) = [p \wedge \neg q''](w) = 0$. If $q(w) = 0$, we have $[p \wedge q'](w) = [p \wedge q''](w) = 0$ and $[p \wedge \neg q'](w) = [p \wedge \neg q''](w) = 1$.

Then, $p \wedge q' = p \wedge q''$ and $p \wedge \neg q' = p \wedge \neg q''$, from which the fact that $F_p(q') = F_p(q'')$ immediately follows. Therefore $F_p^D(q)$ is defined.

As for the other direction, the same reasoning as in the previous result shows that if there is w such that $q(w) = \#$ and $p(w) = 1$, then $F_p^D(q)$ is not defined. By contraposition, if $F_p^D(q)$ is defined, there is no such w. □

Result 14 (Deployment of disjunction)**.** For total propositions p and q, define $F_p(q)$ to represent the Karttunen denotation of a disjunctive question:

$$F_p(q) := \lambda w. \lambda r_{st}. r \in \{p, q\} \wedge r(w)$$

Let p be a total proposition that is true in at least two worlds. $F_p^D(q)$ is defined if and only if q is total.

Proof. It is clear that if q is total, $F_p^D(q)$ is defined.

Assume that there is w such that $q(w) = \#$. Take $q' \in q^R$ such that $q(w) = 1$, and $q'' \in q^R$ that is the same as q', except that $q''(w) = 0$.

- If $p(w) = 0$, we have $F_p(q')(w) = \{q'\}$ and $F_p(q'')(w) = \varnothing$.
- If $p(w) = 1$, we have $F_p(q')(w) = \{p, q'\}$ and $F_p(q'')(w) = \{p\}$. If $q' \neq p$, $F_p(q')(w) \neq F_p(q')(w)$. If $q' = p$, take w' such that $p(w') = 1$. We have $F_p^D(q')(w') = \{p, q'\} = \{p\}$ and $F_p(q'')(w') = \{p, q''\}$ (recall that q'' agrees with q', and therefore with p, at w'). Since $q'' \neq q'$ by construction, $F_p(q')(w') \neq F_p(q'')(w')$.

Either way, $F_p^D(q)$ is not defined. By contraposition, if $F_p^D(q)$ is defined, q is total. □

[55]This can also be stated as "p entails $\pi(q)$", where π is defined as in (47).

References

Abenina-Adar, Maayan and Yael Sharvit (2021). “On the presuppositional strength of interrogative clauses”. *Natural Language Semantics* 29.1, pp. 47–90.

Abruśan, Marta (2011). “Presuppositional and negative islands. A semantic account”. *Natural Language Semantics* 19.3, pp. 257–321.

Beaver, David and Emiel Krahmer (2001). “A partial account of presupposition projection”. *Journal of Logic, Language and Information* 10.2, p. 147.

Biezma, Maria and Kyle Rawlins (2012). “Responding to alternative and polar questions”. *Linguistics and Philosophy* 35.5, pp. 361–406.

Bolinger, Dwight (1978). “Yes-no questions are not alternative questions”. In: *Questions*. Ed. by Henry Hiż. Springer, pp. 87–105.

Büring, Daniel and Christine Gunlogson (2000). “Aren’t positive and negative polar questions the same?” Ms., UCSC/UCLA.

Champollion, Lucas, Ivano Ciardelli, and Floris Roelofsen (2017). “On questions and presuppositions in typed inquisitive semantics”. Handout, NYU.

Ciardelli, Ivano, Jeroen Groenendijk, and Floris Roelofsen (2013). “Inquisitive Semantics. A New Notion of Meaning”. *Language and Linguistics Compass* 7.9, pp. 459–476.

Ciardelli, Ivano, Jeroen Groenendijk, and Floris Roelofsen (2018). *Inquisitive semantics*. Oxford University Press.

Ciardelli, Ivano, Floris Roelofsen, and Nadine Theiler (2017). “Composing alternatives”. *Linguistics and Philosophy* 40.1, pp. 1–36.

Dayal, Veneeta (1996). *Locality in WH-quantification. Questions and relative clauses in Hindi*. Springer.

Dotlačil, Jakub and Floris Roelofsen (2019). “Dynamic inquisitive semantics: Anaphora and questions”. In: *Proceedings of Sinn und Bedeutung 23*. Ed. by M.Teresa Espinal et al. University of Konstanz, pp. 365–382.

Dotlačil, Jakub and Floris Roelofsen (2020). “A dynamic semantics of single-*wh* and multiple-*wh* questions”. In: *Proceedings of SALT 30*. Ed. by Joseph Rhyne et al. Linguistic Society of America, pp. 376–395.

von Fintel, Kai (1999). “NPI licensing, Strawson entailment, and context dependency”. *Journal of semantics* 16.2, pp. 97–148.

George, Benjamin R. (2014). “Some remarks on certain trivalent accounts of presupposition projection”. *Journal of Applied Non-Classical Logics* 24 (1-2), pp. 86–117.

Groenendijk, Jeroen (1998). “Questions in update semantics”. In: *Formal semantics and pragmatics of dialogue: Proceedings of the 13th Twente Workshop on language technology (Twendial ’98)*. Ed. by Joris Hulstijn and Anton Nijholt. University of Twente, pp. 125–137.

Groenendijk, Jeroen and Martin Stokhof (1984). “Studies on the semantics of questions and the pragmatics of answers”. PhD thesis. University of Amsterdam.

Hamblin, Charles L. (1976). “Questions in Montague English”. In: *Montague grammar*. Ed. by Barbara H. Partee. Academic Press, pp. 247–259.

Heim, Irene (1983). “On the projection problem for presuppositions”. In: *Formal Semantics: The Essential Readings*. Ed. by Paul Portner and Barbara Partee. Blackwell, Oxford, pp. 249–260.

Hirsch, Aron (2017). “An inflexible semantics for cross-categorial operators”. PhD thesis. MIT.

Hirsch, Aron and Bernhard Schwarz (2019). “Singular *which*, mention-some, and variable scope uniqueness”. In: *Proceedings of SALT 29*. Ed. by Katherine Blake et al. Linguistic Society of America, pp. 748–767.

Hoeks, Morwenna and Floris Roelofsen (2019). “Coordinating questions: The scope puzzle”. In: *Proceedings of SALT 29*. Ed. by Katherine Blake et al. Linguistic Society of America, pp. 562–581.

Karttunen, Lauri (1973). “Presuppositions of compound sentences”. *Linguistic Inquiry* 4.2, pp. 169–193.

Karttunen, Lauri (1974). “Presupposition and linguistic context”. *Theoretical Linguistics* 1.1-3, pp. 181–194.

Karttunen, Lauri (1977). “Syntax and semantics of questions”. *Linguistics and philosophy* 1.1, pp. 3–44.

Krifka, Manfred (2001). “For a structured meaning account of questions and answers”. In: *Audiatur vox sapientia. A festschrift for Arnim von Stechow*. Ed. by Caroline Féry and Wolfgang Sternefeld. Akademie Verlag, pp. 287–319.

Li, Haoze (2019). “Deriving short answers from Hamblin sets: A case study of Mandarin *wh*-conditionals”. In: *Proceedings of SALT 29*. Ed. by Katherine Blake et al. Linguistic Society of America, pp. 138–159.

Partee, Barbara and Mats Rooth (1983). “Generalized conjunction and type ambiguity”. In: *Meaning, Use, and Interpretation of Language*. Ed. by Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow. De Gruyter, pp. 361–384.

Peters, Stanley (1979). “A truth-conditional formulation of Karttunen’s account of presupposition”. *Synthese* 40.2, pp. 301–316.

Roelofsen, Floris (2019). “Semantic theories of questions”. In: *Oxford Research Encyclopedia of Linguistics*. Oxford University Press.

Roelofsen, Floris and Donka F. Farkas (2015). “Polarity particle responses as a window onto the interpretation of questions and assertions”. *Language* 91.2, pp. 359–414.

Rooth, Mats (1992). “A theory of focus interpretation”. *Natural Language Semantics* 1.1, pp. 75–116.

Rothschild, Daniel (2011). “Explaining presupposition projection with dynamic semantics”. *Semantics and Pragmatics* 4.3, pp. 1–43.

Schlenker, Philippe (2008). “*Be articulate*. A pragmatic theory of presupposition projection”. *Theoretical Linguistics* 34.3, pp. 157–212.

Schlenker, Philippe (2009). “Local Contexts”. *Semantics and Pragmatics* 2.3, pp. 1–78.

Schwarz, Bernhard and Alexandra Simonenko (2018). “Decomposing universal projection in questions”. In: *Proceedings of Sinn und Bedeutung 22*. Ed. by Uli Sauerland and Stephanie Solt. University of Konstanz, pp. 361–374.

Singh, Raj (2007). “Assertability constraints and absurd assertions”. Ms., MIT.

Sudo, Yasutada (2013). “Biased polar questions in English and Japanese”. In: *Beyond Expressives: Explorations in Use-Conditional Meaning*. Ed. by Daniel Gutzmann and Hans-Martin Gärtner. Brill, pp. 275–295.

Szabolcsi, Anna (1997). “Quantifiers in Pair-List Readings”. In: *Ways of Scope Taking*. Ed. by Anna Szabolcsi. Springer, pp. 311–347.

Szabolcsi, Anna (2016). “Direct vs. indirect disjunction of *wh*-complements, as diagnosed by subordinating complementizers”. Ms., New York University.

Theiler, Nadine (2021). “An epistemic bridge for presupposition projection in questions”. In: *Proceedings of SALT 30*. Ed. by Joseph Rhyne et al. Linguistic Society of America, pp. 252–272.

Uegaki, Wataru (2019). “The semantics of question-embedding predicates”. *Language and Linguistics Compass* 13.1.
Uegaki, Wataru (2020). “The existential/uniqueness presupposition of *wh*-complements projects from the answers”. *Linguistics and Philosophy*.
Xiang, Yimei (2021). “A hybrid categorial approach to question composition”. *Linguistics and Philosophy* 44, pp. 587–647.

Chapter 7

A bilateral dynamic semantics for questions and propositions

The content of this chapter is an unpublished manuscript based on work in progress; it is in some ways derivative of the previous chapter.

Abstract

The central idea of this chapter is to make use of the formal similarities between the treatment of *wh*-indefinites in alternative semantics, and that of regular indefinites in dynamic semantics, to show the potential of dynamic propositions as a representation of questions. The proposed system is bilateral, which allows for greater parallelism between polar questions and constituent questions than is generally assumed. I show that the system is strictly an extension of existing static representations, thus guaranteeing that no analytical possibilities are lost, and then explore two potential benefits of the extra expressivity it gives us: an analysis of short answers as dynamic clauses (without ellipsis), and a definition of novel, *short-circuiting* connectives, letting us account for the parallel patterns of anaphora resolution in coordinated polar questions and coordinated declaratives.

Acknowledgements

I thank Benjamin Spector for discussion.

7.1 Introduction

7.1.1 Dynamic semantics and alternative semantics

Dynamic semantics has been introduced chiefly to analyse presupposition projection (e.g. Heim 1983b) and anaphoric dependencies involving indefinites (e.g. Heim 1983a; Groenendijk and Stokhof 1991), as in (1).

(1) A man walked in. He whistled.

In the framework of dynamic semantics, as exemplified in particular py DPL (Groenendijk and Stokhof 1991), pronouns like the one in the second sentence of (1) are analysed as denoting individual-type variables, as is generally done in static accounts. However, there is no specific, identifiable individual that we can say the pronoun refers to. The way that DPL solves this conundrum is that it assumes that sentence meanings are non-deterministic (or indeterminate, depending on how one wants to look at it), in the sense that sentences have a range of potential interpretations rather than a single one. What triggers the indeterminacy is the presence of an indefinite, such as *a man* in the above example; dynamic semantics takes indefinites to introduce new variables whose value is left underdetermined.

As pointed out in particular by Charlow (2020), the indeterminacy introduced in DPL is very similar to what has been proposed in alternative semantics to analyse questions (Hamblin 1976) and focus structure (Rooth 1992). In Hamblin's theory, it is *wh*-words that introduce indeterminacy, because they denote non-singleton alternative sets. Charlow (2020) shows that the traditional dynamic analysis can be decomposed into alternative semantics plus a dynamic component, and proposes to extend Hamblin's approach to indefinites as an analysis of the exceptional scope of indefinites.

This paper explores the idea of relying on the similarity between dynamic semantics and alternative semantics to develop dynamic semantics into a theory of questions, as alternative semantics was designed to be, and show the potential analytical benefits. The basic idea is that *wh*-words will act as variable introducers (or in other words, dynamic existential quantifiers), exactly like regular indefinites in DPL — in fact we will not distinguish indefinites and *wh*-words.

The idea that indefinites and *wh*-words should be analysed in tandem is well established in the literature. They are similar morphologically across languages (Haspelmath 1997, a.o.), and they play a similar role of introducing some form of indeterminacy, which allows for formal parallelisms. Thus, Karttunen (1977) treats *wh*-words as generalized existential quantifiers, as per Montague's (1973) analysis of the indefinite article. In the other direction, Kratzer and Shimoyama (2002) extend Hamblin's (1976) analysis of *wh*-words as alternative introducers to non-*wh* indefinites, as does Charlow (2020) who we already mentioned. The assumption that all indefinites are alternative introducers is also found in the paradigm of Inquisitive Semantics (Ciardelli, Groenendijk, and Roelofsen 2018). Finally, analyses of *wh*-words as dynamic existentials are offered, among others, by Aloni and van Rooy (2002), Haida (2008), and Dotlačil and Roelofsen (2019) (based on inquisitive semantics).

The theory presented here is a variation on these accounts and is particularly close to that of Aloni and van Rooy (2002), sharing with it the radical identification of *wh*-words to dynamic existentials, without assuming extra machinery specific to questions. We will innovate in making the system bilateral, which will make a uniform account of constituent questions and polar questions possible, as well as allow for a novel account of the connectives, motivated by patterns of anaphora resolution in coordinated questions.

7.1.2 Bilateralism

The system we are going to propose is in the spirit of Dynamic Predicate Logic (DPL, Groenendijk and Stokhof 1991). In a DPL-like system, sentences denote a relation between assignments. Let us denote the type of assignments as i. The extensional type of a proposition can be written as $i \to i \to t$, or equivalently as $i \to \{i\}$, where $\{\cdot\}$ is a type constructor for sets. The corresponding intensional type is $s \to i \to \{i\}$. From this perspective propositions map input indices (pairs of a world and an assignment) to a (possibly empty) set of output assignments. At a given, single input index, a proposition is true if it relates the input to a non-empty set of outputs. Thus truth is identified to "survival" of inputs.

We adopt here a slightly richer output type for sentences: they will pair their output assignments with *truth values*. Thus their type will be $i \to \{(i;t)\}$. This richer type allows us to disentangle truth and survival. If p is a proposition and g is an assignment, then $p(g)$ might include descendants g' of g that made p true (if $(g';1) \in p(g)$)) or false (if $(g';0) \in p(g)$).

Conceptually, the idea is that a truth value of 1 marks "active" possibilities. 0-marked possibilities are "passive" and bound to be eliminated at some point, but we keep them around while building up the sentence because various operators (such as negation) are susceptible to reactivate them. We will refer to the type t member of sentence outputs as the "activeness tag" or just "tag".

An alternate, equivalent way of looking at this system is to consider that sentences have both a positive and a negative interpretation, which we can write as $[\![\cdot]\!]_+$ and $[\![\cdot]\!]_-$ respectively, corresponding to two projections of propositional meanings written as $[\cdot]_+$ and $[\cdot]_-$. The type of each projection or interpretation is $i \to \{i\}$ (essentially the type of DPL propositions) and they relate to the interpretation with tags as follows:

(2) $[p]_+ = \lambda g.\, \{g' \mid (g';1) \in p(g)\}$
$[p]_- = \lambda g.\, \{g' \mid (g';0) \in p(g)\}$

(3) $[\![\varphi]\!]_+ = [[\![\varphi]\!]]_+$
$[\![\varphi]\!]_- = [[\![\varphi]\!]]_-$

Once truth and survival are disconnected, it is possible to define, for instance, an externally dynamic, involutive negation:

(4) $\neg p = \lambda g.\, \{(g';\neg b) \mid (g';b) \in p(g)\}$

This definition of negation essentially flips the positive and negative denotation, so that:

(5) $[\![\text{not}\ \varphi]\!]_+ = [\![\varphi]\!]_-$
$[\![\text{not}\ \varphi]\!]_- = [\![\varphi]\!]_+$

The system we are going to describe here is largely similar to the account of anaphoric reference to indefinites in "bathroom sentences" (such as (6)) and sentences involving double negation (such as (7)) offered by Krahmer and Muskens (1995, henceforth K&M). The terminology of "active" and "passive" assignments is borrowed from them.

(6) Either there is no bathroom in this building, or it is in a funny place.

(7) It is not true that John didn't bring an umbrella. It was purple and it stood in the hallway.

Charlow (2020) also proposes a similar system as part of a compositional dynamic theory of the binding potential of indefinites. In Charlow's theory, bilateralism is not a goal in itself, but rather a consequence of trying to add dynamicity in a modular way on top of a classical Montagovian theory.[1]

In our case, bilateralism will be a necessary feature of an account of the anaphoric potential of *wh*-words and indefinites occurring in polar questions, as well as allow us to define *short-circuiting* connectives, as part of a novel account of question conjunction and disjunction.

7.2 Proposal

7.2.1 The basics

The basic idea of the theory presented here is that we are going to represent questions as dynamic propositions. We make the unusual choice to make no clear formal distinction between assertions and questions: this is because such a distinction is not needed for our purposes. Of course, a full model of discourse pragmatics would certainly need to make this distinction at some level.

As noted, the semantic type of a sentence in our system is $i \to \{(i;t)\}$. We assume that assignment functions are total, but that at the beginning of discourse all referents point to a dummy element $\bot$. This is again to simplify notation. Also to simplify, we assume that $\bot$ can be any type (or if one prefers, that there is actually $\bot_e$, $\bot_t$, etc. depending on the type of each variable). We assume that for any static atomic predicate P, $P(\bot) = 0$. A final assumption is that variables are subject to a novelty condition, so that if an indefinite is indexed with variable x, it is presupposed that accessible input assignments g are such that $g(x) = \bot$.

The interpretation we associate to "atomic" sentences like "John came" is straightforward: we assume that they return their input assignment unmodified, while setting the tag to the classical truth value of the proposition. An example is given in (8).[2]

(8) $[\![\text{John came}]\!]^w = \lambda g.\, \{(g; \mathbf{came}^w(\mathbf{j}))\}$

As is standard in dynamic semantics, we assume that indefinites introduce a variable ranging over individuals that satisfy the restrictor and scope. When no witness exists for the existential statement, instead of having no output as in DPL, the input is left unchanged with the tag set to zero.

(9) $[\![\text{there is a}^x\text{ bathroom in this building}]\!]^w =$

$$\lambda g.\, \begin{cases} \{(g[x \mapsto a]; 1) \mid \mathbf{bathroom}^w(a) \wedge \mathbf{itb}^w(a)\} & \text{if non-empty} \\ \{(g; 0)\} & \text{otherwise} \end{cases}$$

[1] In a recent manuscript, Elliott (2020) proposes to derive the dynamic meaning of the connectives from properties of trivalent logic within a system based on Charlow's; the resulting connectives have similar short-circuiting properties as those we are going to define in section 7.3.2.

[2] Presumably "John came" also adds a referent to John as well as a propositional referent to "John came" to the assignment function; we ignore these possibilities to keep the examples simple.

As far as negative indefinites are concerned, we can simply assume that they are equivalent to negated indefinites:

(10) $[\![\text{there is no}^x \text{ bathroom in this building}]\!]^w =$

$$\lambda g. \begin{cases} \{(g[x \mapsto a]; 0) \mid \mathbf{bathroom}^w(a) \wedge \mathbf{itb}^w(a)\} & \text{if non-empty} \\ \{(g; 1)\} & \text{otherwise} \end{cases}$$

Note that above and below, we use x, y etc. for type e variables in the domain of the assignment function, whereas a, b, c etc.are used for metalanguage variables ranging over individuals.

7.2.2 Question words as variable introducers

The traditional treatment of indefinites in dynamic semantics essentially incorporates a notion of alternatives that is very similar to what is done in alternative semantics and in the answer set family of theories of questions (e.g Hamblin 1976; Karttunen 1977). It is thus possible to build an account of question meaning using the dynamic representation of propositions (this has in fact already been explored by Aloni and van Rooy (2002) among others). This representation is in fact a bit richer than answer set theory, which will allow us to account for dynamic effects involving presupposition projection and anaphora resolution of course, but also to incorporate some insights from categorial theories of questions. As we will see, it also makes a novel account of question coordination possible.

The key idea is, following Aloni and van Rooy (2002), to apply the classical analysis of indefinites within dynamic semantics to *wh*-indefinites. Thus, DP *wh*-words like *who* in (11) are assumed to introduce a new type *e* variable and restrict it to the restrictor and scope, as seen in (12). Thus, in our system, there will be no difference at all between *who came* and *someone came*.[3]

(11) Who came?

(12) $[\![\text{who}^x \text{ came}]\!]^w =$

$$\lambda g. \begin{cases} \{(g'; 1) \mid g'[x]g \wedge \mathbf{came}^w(g'(x))\} & \text{if this set is non-empty} \\ \{(g; 0)\} & \text{otherwise} \end{cases}$$

As far as simple polar questions are concerned, we assume that they are essentially *wh*-questions where the restrictor is a singleton set. We can take the unique value in the restrictor to be an arbitrary dummy value that we write as α. This is not going to be important, and it could as well be, for instance, a propositional referent to the nucleus.[4] Embedded polar questions can be introduced by what looks like a *wh*-word, *whether*; we assume that this is also the case of matrix polar questions, though *whether* can be silent (in which case we write ~~*whether*~~). Thus, if φ is a proposition and u is a variable, then:

[3]We ignore the restriction of *who* to human beings to alleviate notation.

[4]One possibility is to identify the focussed element of a polar question with the indexing variable; e.g. "did JOHN see Mary?" would be equivalent to "who saw Mary?" where *who* ranges exclusively over John, while "did John see MARY?" would be equivalent to "who did John see?" where *who* ranges exclusively over Mary. This syntax of polar questions in Turkish, where the apparent question particle will be attached to various elements depending on context (cf. Kamali and Büring 2011) ties in well with such an approach.

(13) $$[\![\text{whether}^u\ \varphi]\!]^w = \lambda g.\{(g'[u \mapsto \alpha]; 1) \mid g' \in [\![\varphi]\!]^w_+(g)\} \cup \{(g'; 0) \mid g' \in [\![\varphi]\!]^w_-(g)\}$$

Concretely, if we start from an arbitrary assignment, we might have:

(14) $$[\![\text{whether}^u \text{ John came}]\!]^w(g) = \begin{cases} \{(g[u \mapsto \alpha]; 1)\} & \text{if } \mathbf{came}^w(\mathbf{j}) = 1 \\ \{(g; 0)\} & \text{otherwise} \end{cases}$$

Finally, multiple-*wh* questions can also be accounted for as questions with a vector of *wh*-variables:

(15) $$[\![\text{who}^x \text{ danced with who}^y]\!]^w(g) = \begin{cases} \{(g'; 1) \mid g'[x, y]g \wedge \mathbf{danced\text{-}with}^w(g'(x), g'(y))\} & \text{if this set is nonempty} \\ \{(g; 0)\} & \text{otherwise} \end{cases}$$

Our main purpose here is to discuss question representations, and to this end we do not need to provide a full discourse model. It might be helpful however to give a sketch of what such a model might look like. Following dynamic tradition, the discourse context should include a set of accessible indices, in the form of world-assignment pairs. Recall that, as we already noted, there is no formal distinction between questions and declaratives in this system: they are both bilateral dynamic propositions. Nevertheless, it is clear that a declarative and the corresponding polar question should not update the context in the same way. A natural assumption is that both declaratives and interrogatives are evaluated against all accessible indices; for declaratives, we form a new context set by collecting all active possibilities in the output, associated to the corresponding world of evaluation. For interrogatives, in contrast, we collect both active and passive possibilities; thus an interrogative will not eliminate any world.[5]

Additionally, we can note that the main defining property of a question is that a certain variable, or a vector of variables (in the case of multiple-*wh* or coordinated questions), was introduced by a *wh*-word. We will refer to such variables as *wh*-variables. It is necessary to know what the *wh*-variable is in order to understand what the question is about; a full discourse model would therefore need to feature a list of currently active *wh*-variables (cf. Aloni and van Rooy 2002 for an example of a system with similar properties). New variables should be added to the list when a question is asked, and old variables should be removed when their value is settled throughout the context. We can partition the accessible indices by grouping them with respect to the values they assign to the *wh*-variables; this partition can serve as a notion of QUD relative to which contextual relevance can be defined, in the spirit of Roberts (1996).

7.2.3 Reduction to static theories

In order to show that the theory proposed here does not force us to abandon previous analyses, the simplest way is to show how we can reduce our representation of questions to

[5] Here we do not consider the possibility that a proposition should output no possibilities, active or passive, at a world; possibly, allowing for this case could be a way of representing presupposition failure in the system.

One alternative to having a different update rule for interrogatives and declaratives would be to assume that all interrogatives are headed by an operator that makes all output possibilities active. This operator should be able to take scope above connectives, if the account of question coordination offered in section 7.3.2 is to be maintained.

established static theories.

Chierchia (2000) remarks that, in dynamic theories, existentials (analysed as variable introducers) and predicates are essentially interchangeable. Indeed, the operation of *existential disclosure* lets us extract the inner predicate from an existential statement. The key insight is that, in DPL and similar systems, (16a) and (16b) are equivalent.

(16) a. $\lambda a. [\exists x. P(x)] \wedge a = x$
b. $\lambda a. P(a)$

We can apply this insight to our system and define a disclosure operator $\{\cdot\}^{\square}$ that turns an intensional dynamic proposition into an intensional predicate, by applying existential disclosure relative to a certain variable written as a superscript; the variable in question will of course have to be the *wh*-variable.[6]

(17) **Abstract formation.** For p with type $s \to i \to \{(i;t)\}$:
$\{p\}^x_g = \lambda a. \lambda w. \exists v \in \{0,1\}. \exists g'. (g';v) \in p(w)(g) \wedge g'(x) = a$

If we look at the output of abstract formation as a function from individuals to propositions, we see that when applied to an individual a other than the empty one ($\bot$), it will return the proposition that a is a witness for the question; for instance, for (11), it will be the proposition that a came. Indeed, the worlds where the *wh*-variable is set to a in an output are exactly those where a is a witness. Thus, ignoring the case of $\bot$, the output of abstract formation is similar to the categorial denotations for questions proposed for instance by Xiang (2021), in turn derivative of the "question abstracts" of Groenendijk and Stokhof (1984), hence the name. One unusual feature, however, is the output of a negative proposition when the first argument is $\bot$: for (11) it will be the proposition that nobody came. Since we assume that the input assignment is unmodified in worlds where no witness exists, and that the *wh*-variable is subject to a novelty condition (so that the input assignment g will always set it to $\bot$), then the output will set the *wh*-variable to $\bot$ in exactly those worlds where there is no witness. All this is seen in the examples below.

(18) a. $\overbrace{\{[\![\text{who}^x \text{ came}]\!]\}^x_g}^{Q} = \lambda a. \lambda w. \begin{cases} \textbf{came}^w(a) & \text{if } a \neq \bot \\ \neg\exists b. \textbf{came}^w(b) & \text{if } a = \bot \end{cases}$

b. Informally:
$Q([\![\text{John}]\!]) = \lambda w. \text{John came in } w$
$Q(\bot) = \lambda w. \text{nobody came in } w$

(19) $\{[\![\text{who}^x \text{ danced with who}^y]\!]\}^{x,y}_g =$
$\lambda a, b. \lambda w. \begin{cases} \textbf{danced-with}^w(a,b) & \text{if } a, b \neq \bot, \bot \\ \neg\exists c, d. \textbf{danced-with}^w(c,d) & \text{otherwise} \end{cases}$

When it comes to polar questions, there are only two possible values for the first argument, the dummy value α and the empty value $\bot$. The abstract will respectively return the positive and negative answers to the question.

[6] In principle we could and should keep track of the dependency to the input assignment in the same way as the dependency to the world; we assume the assignment is fixed here for the sake of simplicity.

(20) a. $\overbrace{\{[\![\text{whether}^u \text{ John came}]\!]\}^u_g}^{Q'} = \lambda u.\, \lambda w.\, [u = \alpha \wedge \mathbf{came}^w(\mathbf{j})] \vee [u = \bot \wedge \neg \mathbf{came}^w(\mathbf{j})]$

b. $Q'(\alpha) = \lambda w.$ John came in w

$Q'(\bot) = \lambda w.$ John did not come in w

In these examples, we can see that when the first argument of the abstract is saturated, the propositions we obtain are answers to the questions in the sense of the answer set theory. Thus, if we let the first argument range over its domain and collect the results in a set, we can derive answer sets from the abstracts, and thus end up with denotations of the sort proposed by Karttunen (1977) or by Hamblin (1976); The operator ALT below, which applies again to dynamic propositional intensions, derives Karttunen denotations (mapping worlds to sets of true propositions), using abstracts as an intermediate result — Hamblin denotations can be obtained by taking the union of all possible Karttunen denotations. Like the abstract formation operator, ALT takes a variable name as argument which corresponds to the *wh*-variable.[7] As we can see, we recover roughly the usual denotations of questions in answer set theory, but for the addition of a negative answer.

(21) **Answer set.** For p with type $s \to i \to \{(i;t)\}$:
$\text{ALT}^x_g(p) = \lambda w.\ \big\{\{p\}^x_g(a) \mid \{p\}^x_g(a)(w)\big\}$

Whence:

(22) $\text{ALT}^x_g\,([\![\text{who}^x \text{ came}]\!]) =$

$$\lambda w.\ \begin{cases} \{\lambda w'.\, \mathbf{came}^{w'}(a) \mid \mathbf{came}^{w}(a)\} & \text{if someone came in } w \\ \{\lambda w'.\, \neg\exists a.\, \mathbf{came}^{w'}(a)\} & \text{otherwise} \end{cases}$$

(23) $\text{ALT}^{x,y}_g\,([\![\text{who}^x \text{ danced with who}^y]\!]) =$

$$\lambda w.\ \begin{cases} \Big\{\lambda w'.\, \mathbf{danced\text{-}with}^{w'}(a,b) \mid \mathbf{danced\text{-}with}^{w}(a,b)\Big\} & \text{if two people danced in } w \\ \{\lambda w'.\, \neg\exists a,b.\, \mathbf{danced\text{-}with}^{w}(a,b)\} & \text{otherwise} \end{cases}$$

(24) $$\text{ALT}^u_g\,([\![\text{whether}^u \text{ John came}]\!]) = \lambda w.\ \begin{cases} \{\lambda w'.\, \mathbf{came}^{w'}(\mathbf{j})\} & \text{if John came in } w \\ \{\lambda w'.\, \neg\mathbf{came}^{w'}(\mathbf{j})\} & \text{if John did not come in } w \end{cases}$$

Relative to typical accounts, there are two oddities here. First, there is an extra negative case in constituent questions. It is common to assume with Dayal (1996) among others that constituent questions presuppose the existence of a witness, and to omit the negative answer in the answer set. The evidence for a general exstential presupposition is in my view rather weak, given that the dialogue in (25) is completely natural. Furthermore, including this extra answer makes for greater parallelism between constituent questions and polar questions. Nevertheless, our formalism makes it easy to allow constituent questions not to have a negative answer in their answer set; a simple approach consists in eliminating passive possibilities from the dynamic denotation of constituent questions, for instance by means of a silent operator. Applying $\{\cdot\}$ or ALT would then lead to a classical answer set with only positive answers. For polar questions, this operation would lead to a trivial answer set with only one possibility.

[7] Also like the abstract formation operator, ALT could and should return sets of assignment-dependent static propositions with type $s \to i \to t$, rather than rely on a single assignment. As before, this choice is to simplify the exposition.

(25) Q: Who came?
A: Nobody.

The second oddity has to do with the condition for membership in the set. Where we depart from both categorial and answer set-based approaches is in that we do not directly represent a domain for the question: it is not possible with this representation to separate the domain of the *wh*-word (here, human beings) and the scope (here, individuals who came). This will limit our expressivity in contexts where the domain is not contextually fixed. To avoid this limitation, we restrict here our attention to contexts where the domain is contextually fixed.[8] In the case of polar questions, given that the domain is reduced to a dummy element, the matter does not really arise.

Our definitions for both $\{\cdot\}$ and ALT do not take into account the activeness tags in any way; thus there is no distinction between positive and negative answers in the output. This is consistent with common assumptions with both categorial and answer set-based approaches, where no such distinction is made. However, it is also possible to map our question denotations to trivalent inquisitive predicates in the style of Enguehard's (2021), using the operator BIALT defined in (26); thus the theory can also be seen as an extension of Enguehard's.

(26) **Trivalent inquisitive denotation.** For p with type $s \to i \to \{(i; t)\}$:

$$\textsc{BiAlt}^x_g(p) = \lambda s. \begin{cases} 1 & \text{if } \exists a. \forall w \in s. \exists g' \in [p(w)]_+(g). g'(x) = a \\ 0 & \text{if } \exists a. \forall w \in s. \exists g' \in [p(w)]_-(g). g'(x) = a \\ \# & \text{in all other cases} \end{cases}$$

Our interim conclusion is that the dynamic theory of questions we propose here is expressive enough to integrate any analysis based on other popular approaches, including categorial theories, and Hamblin-Karttunen semantics. It also offers an analysis of polar questions that is more closely parallel to its analysis of constituent questions than is generally proposed. The next section demonstrates a few benefits of the extra expressivity.

7.3 Applications

7.3.1 Short answers and nested short answers

Constituent questions can be answered by simply supplying the missing element, as in (27). Such answers are known as short answers.

(27) Q: Who came?
A: John.

In answer set semantics, the set of potential short answers cannot be extracted from the question denotation. Instead, short answers are analysed as elliptic forms of full clauses corresponding to members of the answer set. In categorial theories, however, short answers can be analysed as regular DPs that combine with the functional question denotation. Jacobson (2016) defends such an analysis on the basis, among other arguments, that it is necessary to account for the contrast shown in (28). Using a short answer as in (28a) is only possible

[8] This is certainly a disadvantage of the current system; I believe slight modifications should allow to get around this limitation.

if the referent (Jill) is known to be part of the domain of the question (mathematics professors). There is no such restriction for a long-form answer as in (28b); in fact, the answer in (28b) is perhaps a bit more natural if Jill is not a mathematics professor, possibly due to competition with the short answer. Another contrast is that a short answer is obligatorily interpreted exhaustively, while for a long answer this is optional (see Jacobson 2016 for details).

(28) Context: *Jill is a physics student.*
Q: Which mathematics professor left the party at midnight?
a. #A: Jill.
b. A: Jill left the party at midnight.

(adapted from Jacobson 2016)

Such a contrast is difficult to account for based on answer set theory, where answerhood is a property of propositions, and (28a) and (28b) are assumed to denote the same proposition.

Since our theory can be reduced to a categorial theory, it is essentially compatible with Jacobson's analysis of question-answer pairs as a syntactic construction, where the question, interpreted as an abstract, takes the answer as an argument, or vice-versa if the answer is a quantifier. We can also however stay at the dynamic level, and analyse the relation between short answers and questions as a semantic one, mediated through binding. The key assumption is that short answers bear an index corresponding to the *wh*-variable; this allows us to convert the type *e* element (or whatever it is) to a proposition that will be interpretable as an answer in the context of the question. We implement this through an indexed operator SHORT. SHORT can combine with any type of answer; its type is $a \to t$ where a is the type of the argument variable. Because of a technical limitation of the system that we will come back to shortly, SHORT returns a mere truth value, that is, we are going to interpret short answers as static propositions rather than dynamic ones. Additionally, its evaluation is parameterized by a world and a *set* of assignments (type $\{i\}$), rather than a world and an assignment. The definition of SHORT is given in (29).

(29) $[\![\text{SHORT}^x]\!]^{w,G} = \lambda a.\, \exists g \in G.\, g(x) = a$

Recall the following assumptions from the informal discourse model introduced in section 7.2.2: at any given point in the discourse, the context makes a certain set of indices, corresponding to world-assignment pairs, accessible. When a sentence is uttered, a new context is obtained by applying its interpretation to all indices and collecting the output assignments, associated to their world, in a grand union; for questions we keep both active and passive possibilities, while for assertions we keep only active ones. This system is not complete — in particular, we are not keeping track of which variables are *wh*-variables — but it is sufficient to illustrate the interpretation of short answers. We will consider that the short answer is evaluated relative to the actual world, as well as the set of all assignments associated to that world in the post-question context.

Assume then that "Who^x came?" is uttered in some context. The output context set will pair all worlds where Mary came, and exactly those worlds, with an assignment where x refers to Mary; this follows from the interpretation showcased in (12). Thus, over the world – set of assignments pairs present in the post-question context, "SHORT^x Mary" will be true exactly in the cases where Mary came in the world being considered. The short answer therefore has the same truth conditions as the declarative sentence “Mary came”.[9]

The analysis can extend to the case of complex quantifiers used as short answers, as in (30), if we simply assume that they take scope above SHORT, as in for example (31). Here the reason we deviated from the system in the definition of SHORT can be given: if we had defined SHORT to create normal dynamic propositions, it would be impossible to give a meaning for *everyone* that would lead to appropriate truth conditions for (31b). This is is because dynamic propositions only see one input assignment at a time, and *everyone* as a short answer needs to see all assignments at a given world at once.[10] We could overcome this limitation of the system, and offer an account of complex short answers as proper dynamic propositions, by replacing assignment functions with sets of assignment functions everywhere, as in systems of dynamic semantics designed to capture plural dependencies (van den Berg 1996; Brasoveanu 2008; Brasoveanu 2010). I leave the integration of the present system and plural semantics to future work.

(30) Who came?
  a. John or Mary.
  b. Everyone.

(31) a. [LIFT[John] or LIFT[Mary]] [$\lambda y.$ SHORT^x t_y]
  b. Everyone [$\lambda y.$ SHORT^x t_y]

In this perspective, the natural analysis of the answer particles “Yes” and “No” is that they denote the dummy particle α and the $\perp$ value of the same type respectively.

Finally, our approach to short answers extends gracefully to multiple-*wh* questions. (32a) below can be analysed straightforwardly as (32b).

(32) a. Q: Who danced with who?
  A: Mary, with John.
  b. Q: Who^x danced with who^y?
  A: SHORT^{x,y} [Mary, John]

An interesting property of the system is that, given the absence of a distinction between *wh*-word and regular indefinites, it is possible to reinterpret questions containing regular indefinites as more complex questions. This lets us in particular analyse cases where extra material is present in the short answer, calling back to an indefinite, as in (33) and (34).[11]

---

[9] Here no attempt is made to derive an exhaustive interpretation; this in fact desirable for two reasons. First, short answers to mention-some questions are not interpreted exhaustively, as seen below.

(i) Q: Where can we watch the game?
  A: At the corner pub.
  $\not\rightsquigarrow$ We cannot watch it anywhere else.

Second, when short answers are in the scope of a truth-conditional operator, the meaning we observe is what we expect if short answers have non-exhaustive semantics, as seen below. This will fall out naturally if we derive the exhaustive interpretation through wide-scope exhaustification in the usual way, rather than building it in.

(ii) Q: Who came?
  A: Not Mary.
  $\rightsquigarrow$ Mary did not come (stronger than: Mary did not come or other people came).

[10] I thank Yasutada Sudo for alerting me to this issue.

[11] However, we cannot account for cases where this sort of sprouting short answers attach to negative an-

(33) a. Q: Did anyone come?
A: Yes, John.
b. Potential analysis:
Q: ~~whether~~^{u} anyone^x came?
A: SHORT^{u,x} [Yes, John]

(34) a. Q: Which girl danced with someone?
A: Mary, with John.
b. Potential analysis:
Q: Which^x girl danced with someone^y?
A: SHORT^{x,y} [Mary, John]

7.3.2 Short-circuiting connectives and anaphora resolution in coordinated polar questions

The connectives Krahmer and Muskens (1995) solve the problem of "bathroom sentences" in the following way: first, they assume that some quantifiers may introduce passive as well as active referents, and in particular, a negative quantifier like *no* introduces passive referents to any individuals that fit the restrictor and scope.[12] This is what we are also assuming.

The second assumption K&M make is that both conjunction and disjunction are internally dynamic. This is a departure from standard DPL or DRT, where conjunction is dynamic, but disjunction is both internally and externally static. We are going to follow K&M on this point too.

To define the connectives, it will be useful to define an operator for dynamic sequencing, which will write as $\oplus$. Note that what we call dynamic sequencing is simply DPL conjunction (written $\wedge$ in DPL). It is also sometimes written with a semicolon (;) as part of an analogy with programming languages.

(35) **Dynamic sequencing.** For p and q with type $i \to \{i\}$ (DPL propositions):
$p \oplus q = \lambda g. \{g'' \mid \exists g'. g' \in p(g) \wedge g'' \in q(g')\}$

We can now write down our internally dynamic analysis of disjunction and conjunction. These connectives can be called *short-circuiting*:[13] they do not feed all of the left-hand output to the right-hand expression.

swers, such as (i).

(i) Q: Did any professor come?
A: No, only John (who is a student).

[12]Note: all quantifiers seem to license plural anaphors to their restrictor set (e.g., the set of bathrooms in (6)), regardless of the precise quantifier and the scope, and regardless of the position of the anaphor relative to the quantifier. While this could be represented in our system by having two index variables on each quantifier, we are ignoring it for now, and focussing on the liceity of reference to the scope set (the bathrooms in this building).

[13]The term *short-circuiting* is borrowed from programming. In most programming languages, the logical connectives behave exactly as in our analysis: the right-hand side is only evaluated if the left-hand side does not already determine the result. For instance, in C, if `arr` is an array, and `i` is an integer variable whose value is negative, attempting to evaluate `arr[i]` is likely to make the program crash. However, evaluating the boolean expression `i >= 0 && arr[i] == 2` will never trigger this error, because the second conjunct will only be evaluated in the event the first one evaluates to 1, i.e. when the value of `i` is nonnegative.

(36) **Disjunction.** For p and q with type $i \to \{(i;t)\}$ (propositional extensions):
$p|q = \lambda g.\{(g';1) \mid (g';1) \in p(g)\} \cup \{(g'';v) \mid \exists g'.(g';0) \in p(g) \wedge (g'';v) \in q(g')\}$

(37) **Conjunction.** For p and q with type $i \to \{(i;t)\}$ (propositional extensions):
$p\&q = \lambda g.\{(g';0) \mid (g';0) \in p(g)\} \cup \{(g'';v) \mid \exists g'.(g';1) \in p(g) \wedge (g'';v) \in q(g')\}$

The logic of these definitions becomes clearer when we look at the active and passive projections and make use of the sequencing operator that we have just defined:

(38)
$$\begin{aligned} [p|q]_+ &= [p]_+ \cup ([p]_- \oplus [q]_+) \\ [p|q]_- &= [p]_- \oplus [q]_- \\ [p\&q]_+ &= [p]_+ \oplus [q]_+ \\ [p\&q]_- &= [p]_- \cup ([p]_+ \oplus [q]_-) \end{aligned}$$

Similar, short-circuiting definitions are already proposed by K&M, except that they make disjunctions and negated conjunctions externally static, while ours are externally dynamic.[14]

In what follows, we overload the symbols & and | to also be defined on propositional intensions, in the obvious way.

Conjunctive questions As discussed by Enguehard (2021), presupposition satisfaction exhibits similar patterns in coordinated polar questions as in assertions. This observation can be extended to anaphora; Groenendijk (1998) remarks on the following example:

(39) Did you see a man? and was he angry? (Groenendijk 1998)

Note that (39) might presuppose that you saw at most one man; if you saw several, and only some of them are angry, it is not actually clear how to answer the question. To keep the discussion simple, I am not going to use Groenendijk's example, and focus on cases where it is plausibly granted that the referent is unique if it exists, so that the fact singular pronouns are felicitous is not remarkable.[15] This keeps us from engaging with the debate on whether donkey pronouns referring back to indefinites should receive universal or existential readings.[16]

As seen below, the pattern we can observe is that when two polar questions are conjoined, whether at matrix or at embedded level, an anaphor in the second conjunct may be bound by an indefinite in the first conjunct; however, no binding is possible if the indefinite is negated.

(40) a. Is there a bathroom, and is it clean?
b. #Is there no bathroom? and is it clean?

[14]While I leave out a detailed analysis, the following examples provide justification for this choice:

(i) Does he have a cat↗ or a dog?↗ and what color is it?

(ii) Is there no cat and no dog in this house? or is it hiding ?

[15]"Bathroom sentences" are in this category, even though they logically should not be: it is of course possible that there might be several relevant bathrooms, and yet it is always felicitous to speak of "the bathroom" regardless of how many bathrooms they are, suggesting that we are always ready to accommodate a contextual restriction such that at most one bathroom exists.

[16]Note that K&M adopt a more sophisticated definition of disjunction than we have here because they want to generate a universal reading.

This is surprising under a "yes/no-symmetric" view of polar questions, such as the (static) Hamblin/Karttunen view that "whether p" is more or less $\{p, \neg p\}$ and the partition theory of Groenendijk and Stokhof (1984), as well as dynamic theories based upon it (e.g. Groenendijk 1998). Such views make the first conjuncts of (40a) and (40b) equivalent, and we therefore expect both examples to have the same status (presumably, to both be infelicitous).

Some other dynamic approaches essentially integrate a static theory of questions with a dynamic view of propositional content: this is the case of the Dynamicised Hamblin Sets of Li (2019) and of the Dynamic Inquisitive Semantics of Dotlačil and Roelofsen (2019). In both cases, yes/no-symmetry is avoided, but since the asymmetry does not affect the workings of conjunction, the prediction that (40a) should be infelicitous is imported from the underlying static theory.

In the system presented above, the paradigm we observe is predicted as a consequence of the short-circuiting nature of conjunction. Let us call p the propositional intension corresponding to "whether^u there is a^x bathroom", and q the propositional intension corresponding to "whether^v it^x is clean." Imagine there are three worlds under consideration, $w_\varnothing$ where there is no bathroom, w_d where there is a dirty bathroom $\mathbf{x}_d$, and w_c where there is a clean bathroom $\mathbf{x}_c$. Then:

(41)
a. $(p\&q)(w_\varnothing)(g) = \{(g; 0)\}$

b. $(p\&q)(w_d)(g) = q(w_d)\left(g\begin{bmatrix} u \mapsto \alpha \\ x \mapsto \mathbf{x}_d \end{bmatrix}\right) = \left\{\left(g\begin{bmatrix} u \mapsto \alpha \\ x \mapsto \mathbf{x}_d \end{bmatrix}; 0\right)\right\}$

c. $(p\&q)(w_c)(g) = q(w_c)\left(g\begin{bmatrix} u \mapsto \alpha \\ x \mapsto \mathbf{x}_d \end{bmatrix}\right) = \left\{\left(g\begin{bmatrix} u \mapsto \alpha \\ v \mapsto \alpha \\ x \mapsto \mathbf{x}_c \end{bmatrix}; 1\right)\right\}$

At no point do we try accessing an unset referent. We can verify that the Karttunen denotation ends up as follows:

(42)
a. $\text{ALT}_g^{u,v}(p\&q)(w_\varnothing) = \{\{p\&q\}_g^{u,v}(\bot, \bot)\}$
$= \{\lambda w.\,\text{there is no bathroom in } w\}$

b. $\text{ALT}_g^{u,v}(p\&q)(w_d) = \{\{p\&q\}_g^{u,v}(\alpha, \bot)\}$
$= \{\lambda w.\,\text{there is a bathroom that is not clean in } w\}$

c. $\text{ALT}_g^{u,v}(p\&q)(w_c) = \{\{p\&q\}_g^{u,v}(\alpha, \alpha)\}$
$= \{\lambda w.\,\text{there is a bathroom that is clean in } w\}$

Thus the partition corresponding to the question has three cells, corresponding to the propositions that there is no bathroom ($\neg p$), that there is an unclean bathroom ($p \wedge \neg q$) and that there is a clean bathroom ($p \wedge q$). This is indeed the issue that (40a) raises. Enguehard (2021) defends this three-way partition as a good analysis of conjunctive questions in general. Nevertheless, the system presented here can also generate the more traditional analysis as a four-way partition ($\{p \wedge \neg q, p \wedge q, \neg p \wedge \neg q, \neg p \wedge q\}$); for this, it is sufficient to assume that an optional operator can set all the tags in the first conjunct's output to 1. In this case, however, anaphora refrring back to the first conjunct should not be possible.

Disjunctive questions Similar patterns obtain in the case of disjunctive questions, thus replicating the "bathroom" data Krahmer and Muskens (1995) are concerned with. Whether the disjunctive question receives the intonation typical of "open" or "closed" readings (the latter a.k.a. "alternative", cf. Roelofsen and Farkas 2015 for the terminology) does not make a difference.

(43) a. Is there no bathroom,↗ or is it dirty?↗
b. Is there no bathroom,↗ or is it dirty?↘
c. #Is there a bathroom, or is it dirty?

Once again, we observe the same patterns as in assertions. This is surprising under a yes/no-symmetric view of polar questions, or from the classical dynamic perspective of disjunction being internally static. The system described here does predict this pattern. Recall that we have three representative worlds $w_{\varnothing}$ (no bathroom), w_d (dirty bathroom) and w_c (clean bathroom). We take p to stand for the proposition denoted by "whether^u there is no^x bathroom" and q for the proposition denoted by "whether^v it^x is dirty".

(44) a. $(p|q)(w_{\varnothing})(g) = \{(g[u \mapsto \alpha]; 1)\}$
b. $(p|q)(w_d)(g) = q(w_d)\,(g\,[x \mapsto \mathbf{x}_d]) = \left\{\left(g\begin{bmatrix} v \mapsto \alpha \\ x \mapsto \mathbf{x}_d \end{bmatrix}; 1\right)\right\}$
c. $(p|q)(w_c)(g) = q(w_c)\,(g\,[x \mapsto \mathbf{x}_c]) = \{(g\,[x \mapsto \mathbf{x}_c]\,; 0)\}$

This time, the members of the answer set will be the proposition that there is no bathroom, the proposition that there is a clean bathroom, and the proposition that there is a dirty bathroom: the partition can be schematized as $\{p, \neg p \wedge q, \neg p \wedge \neg q\}$. These are indeed the good answers to the open question in (43a). The extra presuppositions of the closed reading are not accounted for for now.

Relation of the connectives to other theories Beyond the issue of anaphora resolution, the connectives we defined act on answerhood conditions in a certain way, and we can try to compare it to what is done in H/K semantics. The dominant assumptions for answer set connectives are that conjunction is pointwise propositional conjunction, while disjunction is simply set union. The following is a formal definition in the case of Karttunen denotations:

(45) a. $Q \wedge Q' = \lambda w. \{p \wedge q \mid p \in Q(w) \wedge q \in Q'(w)\}$
b. $Q \vee Q' = \lambda w. Q(w) \cup Q'(w)$

The action of the short-circuiting connectives is not equivalent to the definitions in (45), and in fact it is impossible to describe in terms of the answer set, because the combinatorics make use of the value of the tags; the definitions of $\{\cdot\}$ and ALT do not take the tags into account and therefore, the distinction between positive and negative possibilities is lost at the level of the answer set. What the short-circuiting connectives do to the alternatives is however equivalent to the action of the trivalent connectives of Enguehard (2021).

It is nevertheless possible to reconstruct Hamblin-Karttunen analyses of question coordination within the present system, if one so wishes. We have already mentioned the possibility of introducing an operator that sets all the tags to 1:

(46) **Positivity operator.**[17]For p with type $i \to \{(i; t)\}$:
$\textsc{pos}(p) = \lambda g. \{(g'; 1) \mid (g'; 1) \in p(g) \vee (g'; 0) \in p(g)\}$

Equivalently: $[\textsc{pos}(p)]_+ = [p]_+ \cup [p]_-$
$[\textsc{pos}(p)]_- = \varnothing$

It can be verified that if p and q are two questions with no free variables, whose *wh*-variables are x and y respectively, then $\textsc{alt}_g^{x,y}(\textsc{pos}(p)\&q)$ is the pointwise conjunction of $\textsc{alt}_g^x(p)$ and $\textsc{alt}_g^y(q)$; in other words, taking the short-circuiting conjunction of $\textsc{pos}(p)$ and q leads to the same answer set as applying the conjunction defined in (45) to the answer sets of p and q. This shows we can emulate conjunction. Disjunction as set union can be reconstructed in a similar way, if we allow for an operator that adds a passive referent to every world:

(47) **Passive completion operator.**[18]For p with type $i \rightarrow \{(i;t)\}$:
$\textsc{pass}(p) = \lambda g.\, p(g) \cup \{(g;0)\}$

This time, it can be verified that if p and q are two questions with no free variables, whose *wh*-variables are x and y respectively, then $\textsc{alt}_g^{x,y}(\textsc{pass}(p)|q)$ is the union of $\textsc{alt}_g^x(p)$ and $\textsc{alt}_g^y(q)$.

As far as categorial theories are concerned, they generally do not allow for direct coordination of questions. Indeed, given that categorial questions are functions, with possibly different argument types, natural attempts at defining connectives will not be well-defined. Instead, proponents of categorial theories resort to type-shifting techniques to analyse question coordination (e.g. Krifka 2001; Xiang 2021). Nevertheless, it is interesting to spell out the notions of conjunction and disjunction that our connectives induce on question abstracts. If we assume that the *wh*-variable is set to $\bot$ in the question output if and only if the tag is set to 0 (which is always the case in simple questions in our system, though not necessarily in coordinated questions), then we can define connectives over abstracts that are equivalent to those we define over questions:

(48) **Abstract conjunction.** For two abstracts F and G (respectively type $d \rightarrow s \rightarrow t$ and $d' \rightarrow s \rightarrow t$ for some d, d'):
$F\&\&G = \lambda a, b.\, \lambda w.\, [a = \bot \wedge F(\bot)(w)] \vee [a \neq \bot \wedge F(a)(w) \wedge G(b)(w)]$

(49) **Abstract disjunction.** For two abstracts F and G (respectively type $d \rightarrow s \rightarrow t$ and $d' \rightarrow s \rightarrow t$ for some d, d'):
$F||G = \lambda a, b.\, \lambda w.\, [a \neq \bot \wedge F(a)(w)] \vee [a = \bot \wedge F(\bot)(w) \wedge G(b)(w)]$

Thus, we could have stated our account of anaphora resolution in a categorial theory, though it would have been a significant departure from any existing view on the connectives.

7.4 Conclusion and perspectives

In this paper, I have tried to highlight the viability of dynamic semantics as a theory of questions. The introduction of bilateralism lets us extend the natural analogy between the views of *wh*-indefinites in alternative semantics and regular indefinites in dynamic semantics to polar questions, and provide a parallel account of both. We have seen that the resulting system lets us reuse the tools of "rich" theories of questions such as categorial theories,

[17] As we have done with the connectives, we can overload this operator to also apply to intensions.
[18] Same remark.

while maintaining the possibility to provide an analyses of question coordination, something these theories struggle with. The system as presented here remains in a rather embryonic state, as we have only shown its ability to model distinctions within questions and a few selected phenomena. Future work will strive to develop the theory and study its interaction with the various topics that have received interest in the literature on question semantics, such as answer exhaustivity, the presuppositions of questions, the mention-some / mention-all distinction and the semantics of embedded questions.

References

Aloni, Maria and Robert van Rooy (2002). "The dynamics of question and focus". In: *Semantics and Linguistic Theory*. Vol. 12, pp. 20–39.

van den Berg, Martin H. (1996). "Some aspects of the internal structure of discourse. The dynamics of nominal anaphora". PhD thesis. University of Amsterdam.

Brasoveanu, Adrian (2008). "Donkey pluralities: plural information states versus non-atomic individuals". *Linguistics and Philosophy* 31, pp. 129–209.

Brasoveanu, Adrian (2010). "Decomposing modal quantification". *Journal of Semantics* 27.4, pp. 437–527.

Charlow, Simon (2020). "Static and dynamic exceptional scope". Ms. NYU. Accessed June 2021 at lingbuzz/004650.

Chierchia, Gennaro (2000). "Chinese conditionals and the theory of conditionals". *Journal of East Asian Linguistics* 9, pp. 1–54.

Ciardelli, Ivano, Jeroen Groenendijk, and Floris Roelofsen (2018). *Inquisitive semantics*. Oxford University Press.

Dayal, Veneeta (1996). *Locality in WH-quantification. Questions and relative clauses in Hindi*. Springer.

Dotlačil, Jakub and Floris Roelofsen (2019). "Dynamic inquisitive semantics: Anaphora and questions". In: *Proceedings of Sinn und Bedeutung 23*. Ed. by M.Teresa Espinal et al. University of Konstanz, pp. 365–382.

Elliott, Patrick D. (2020). "Towards a principled logic of anaphora". Ms. MIT. Accessed June 2021 at lingbuzz/005562.

Enguehard, Émile (2021). "Explaining presupposition projection in (coordinations of) polar questions". *Natural Language Semantics*.

Groenendijk, Jeroen (1998). "Questions in update semantics". In: *Formal semantics and pragmatics of dialogue: Proceedings of the 13th Twente Workshop on language technology (Twendial '98)*. Ed. by Joris Hulstijn and Anton Nijholt. University of Twente, pp. 125–137.

Groenendijk, Jeroen and Martin Stokhof (1984). "Studies on the semantics of questions and the pragmatics of answers". PhD thesis. University of Amsterdam.

Groenendijk, Jeroen and Martin Stokhof (1991). "Dynamic predicate logic". *Linguistics and philosophy* 14, pp. 39–100.

Haida, Andreas (2008). "The indefiniteness and focusing of question words". In: *Semantics and Linguistic Theory*. Vol. 18, pp. 376–393.

Hamblin, Charles L. (1976). "Questions in Montague English". In: *Montague grammar*. Ed. by Barbara H. Partee. Academic Press, pp. 247–259.

Haspelmath, Martin (1997). *Indefinite pronouns*. Clarendon.

Heim, Irene (1983a). “File change semantics and the familiarity theory of definiteness”. In: *Meaning, Use, and Interpretation of Language*, pp. 164–189.

Heim, Irene (1983b). “On the projection problem for presuppositions”. In: *Formal Semantics: The Essential Readings*. Ed. by Paul Portner and Barbara Partee. Blackwell, Oxford, pp. 249–260.

Jacobson, Pauline (2016). “The short answer: Implications for direct compositionality (and vice versa)”. *Language* 92.2, pp. 331–375.

Kamali, Beste and Daniel Büring (2011). “Topics in questions”. In: *Workshop on the Phonological Marking of Focus and Topic: GLOW*. Vol. 34.

Karttunen, Lauri (1977). “Syntax and semantics of questions”. *Linguistics and philosophy* 1.1, pp. 3–44.

Krahmer, Emiel and Reinhard Muskens (1995). “Negation and disjunction in discourse representation theory”. *Journal of Semantics* 12.4, pp. 357–376.

Kratzer, Angelika and Junko Shimoyama (2002). “Indeterminate pronouns: the view from Japanese”. In: *Proceedings of the 3rd Tokyo conference on psycholinguistics*. Ed. by Yukio Otsu, pp. 1–25.

Krifka, Manfred (2001). “For a structured meaning account of questions and answers”. In: *Audiatur vox sapientia. A festschrift for Arnim von Stechow*. Ed. by Caroline Féry and Wolfgang Sternefeld. Akademie Verlag, pp. 287–319.

Li, Haoze (2019). “Deriving short answers from Hamblin sets: A case study of Mandarin *wh*-conditionals”. In: *Proceedings of SALT 29*. Ed. by Katherine Blake et al. Linguistic Society of America, pp. 138–159.

Montague, Richard (1973). “The proper treatment of quantification in ordinary English”. In: *Approaches to Natural Language*. Ed. by Jaakko Hintikka, Julius Moravcsik, and Patrick Suppes. Springer, pp. 221–242.

Roberts, Craige (1996). “Information structure in discourse: towards an integrated formal theory of pragmatics”. *OSU working papers in linguistics* 49, pp. 91–136.

Roelofsen, Floris and Donka F. Farkas (2015). “Polarity particle responses as a window onto the interpretation of questions and assertions”. *Language* 91.2, pp. 359–414.

Rooth, Mats (1992). “A theory of focus interpretation”. *Natural Language Semantics* 1.1, pp. 75–116.

Xiang, Yimei (2021). “A hybrid categorial approach to question composition”. *Linguistics and Philosophy* 44, pp. 587–647.

RÉSUMÉ

Cette thèse réunit plusieurs travaux portant sur une variété de sujets liés à la question du rôle joué par les alternatives dans la détermination du sens, et de leur place dans les théories de l'interface sémantique-pragmatique. Un résumé plus fourni peut être trouvé dans les premières pages du document.

MOTS CLÉS

Pragmatique formelle, sémantique formelle, implicatures, alternatives

ABSTRACT

This thesis collects a number of pieces bearing on various questions relating to the role of alternatives in the determination of meaning and their place in the theoretical landscape of the semantics-pragmatics interface. A longer abstract is found at the beginning of the document.

KEYWORDS

Formal pragmatics, formal semantics, implicatures, alternatives