Attentional Pragmatics:
A new pragmatic approach to exhaustivity*

Matthijs Westera
Leiden University Centre for
Linguistics

Abstract The neo-Gricean approach to exhaustivity is based on the idea that exhaustivity arises when relevant propositions are not *asserted*. This paper presents a new pragmatic approach based on the idea that exhaustivity arises when relevant propositions are not *mentioned*, or more precisely, when the speaker did not intend to draw attention to them. This seemingly subtle shift from information to attention results in different predictions on a range of challenges for the neo-Gricean approach, some of which have been brought up in support of the grammatical approach to exhaustivity. This paper discusses three such challenges: exhaustivity on the hints of a quizmaster, exhaustivity on questions, and exhaustivity without an opinionatedness assumption. The two pragmatic approaches are compared on these puzzles along with the grammatical approach.

Keywords: exhaustivity, quantity, disjunction, attention, questions, opinionatedness

1 Introduction

Exhaustivity is a widespread phenomenon whereby, roughly, hearing one thing can lead one to infer the negation of something stronger, for instance from a disjunction (with focus on the disjuncts and a final falling intonation) the negation of a conjunction:

(1)   ALF or BETH was at the protest. *(Implied: not both)*

This utterance implies that Alf and Beth were not both there, even though under a classical, logical treatment of *or* and *and* this does not follow. Indeed, in natural language too it is possible to add *…and maybe both* without contradiction. This phenomenon is variably known as exhaustivity, scalar or quantity implicature/inference/implication.

For many years the only approach to exhaustivity implicature was the neo-Gricean one, which formalizes the following intuitive explanation: if the speaker had believed the stronger information that both Alf and Beth attended, then they would have asserted that — since the speaker didn't, they must believe the stronger information is false. The last two decades have seen the rise of an alternative, grammatical approach to exhaustivity which, motivated by certain challenges for the neo-Gricean approach, has sought to mostly bypass pragmatics and treat exhaustivity, instead, as the result of a grammatical operation. The present paper takes a different route by developing Attentional Pragmatics, a pragmatic theory that offers new solutions to some of the challenges for the neo-Gricean approach.

The core idea is quite simple, and but a minor deviation from the intuition underlying the neo-Gricean approach:

- **Neo-Gricean approach:** Exhaustivity arises when relevant propositions are not *asserted*.

- **Attentional Pragmatics:** Exhaustivity arises when relevant propositions are not *mentioned*.

The crucial difference is that one can mention an alternative without asserting it. For instance, in uttering a disjunction one mentions the disjuncts without asserting either one; and in asking a question one can mention various propositions without asserting anything at all. Of course the relevant notion of 'mentioning' will have to be more precisely characterized, which is done in Attentional Pragmatics in terms of the notion of *attention*. The heart of Attentional Pragmatics is a set of maxims requiring, roughly, that a reasonable

speaker should, in addition to trying to assert all (and only) relevant propositions they consider true as required by the Gricean maxims, also intend to draw attention to all (and only) relevant propositions they consider possible.

The importance of attention in semantics/pragmatics is quite generally acknowledged. For instance, the interpretation of pronouns is guided in part by considering which individuals are attentionally salient, that is, which are available discourse referents. And the common idea that expressions such as disjunction introduce their disjuncts as alternatives (e.g., Alonso-Ovalle 2006, Ciardelli, Groenendijk & Roelofsen 2009) can also be given an attentional understanding: they draw attention to their disjuncts. Central to the current paper is the idea that attention drawing not merely happens and affects interpretation (as is uncontroversial), but happens *intentionally* — that is, drawing attention to things is a type of communicative intention in its own right, just like sharing information.

The current paper presents and motivates Attentional Pragmatics, and compares it to both the neo-Gricean approach and the grammatical approach with regard to three empirical puzzles:

- Exhaustivity appears on the hints of a quizmaster, which Fox (2014) argues are not subject to the maxim of Quantity on which the neo-Gricean approach relies; I present a novel solution that, while developed with Attentional Pragmatics as a backdrop, also works for the grammatical approach.

- Exhaustivity appears on questions (e.g., Geurts 2011, Biezma & Rawlins 2012). Since questions do not serve to assert a piece of information, they are not subject to the maxim of Quantity on which the neo-Gricean approach relies, and also provide no direct way for grammatical exhaustivity operators to affect inferences. Attentional Pragmatics will offer a mostly uniform account of exhaustivity on questions and assertions, while also explaining some differences (foremost that exhaustivity in the case of questions is commonly considered a *presupposition*).

- Exhaustivity occurs in contexts where the speaker is not assumed to be opinionated (Westera 2014, Buccola & Haida 2019, Dieuleveut, Chemla & Spector 2019), which is a problem for the neo-Gricean approach's reliance on an opinionatedness assumption to strengthen inferences from the maxim of Quantity to obtain exhaustivity — and

in fact some implementations of the grammatical approach rely on a similar assumption. The theory proposed here does not.

Elsewhere (Westera 2017a, 2020a) I argue that the shift from the neo-Gricean approach to Attentional Pragmatics also enables a better pragmatic treatment of Hurford disjunctions, that is, disjunctions where one disjunct entails the other, and there I offer a detailed comparison also between Attentional Pragmatics and the grammatical approach in this regard. This huge and important topic had to be left out from the present paper for reasons of scope.

2 Background

2.1 Gricean pragmatics

2.1.1 Content and intent

The approach taken in this paper follows Grice (1975) in trying to derive inferences about utterances by reasoning on the basis of the presumed rationality/cooperativity of the speaker. It is important in this approach to distinguish the sum of all information provided by an utterance from information which the speaker *intended* to provide (i.e., part of what the speaker meant), henceforth an *informational intent*, and that in turn from the information which the utterance expressed by virtue of its semantics alone, henceforth *informational content*.[1] The intent/content terminology is adopted from Bach & Harnish 1979, essentially a rewording of Grice's (1989) "utterer's meaning"/ "sentence meaning". It has the advantage of generalizing easily to attention: *attentional intent* (the set of things to which the speaker intended to draw attention) and *attentional content* (the set of things to which attention was drawn by virtue of the semantics alone). Consistently distinguishing the two notions will help clarify what Attentional Pragmatics does and does not commit us to.

Another distinction on which I will occasionally rely is that between primary and secondary intents, for instance, between the main assertion communicated by an utterance and more backgrounded intents such as conventional and conversational implicatures. This distinction is orthogonal to the

---

1 The distinction between just any information provided and an informational intent is often overlooked in the literature on conversational implicature, where "implicature" is often used interchangeably with "inference" and "implication" (as noted in Bach 2006b) — but an utterance typically implies many things that are not part of what the speaker meant.

informational/attentional distinction; for instance, just as there are primary and secondary informational intents (e.g., the main assertion vs. an implicature), there can be primary and secondary attentional intents (e.g., the main things to which attention is drawn vs. things to which attention is drawn only indirectly or by means of a subordinate clause).

The core of a standard Gricean pragmatic theory is a set of conversational maxims that govern which informational intents (whether primary or secondary, direct or indirect) are rational/cooperative for a given utterance, given the speaker's beliefs about what is true and relevant. The Gricean maxims require, in a nutshell, that informational intents be true (Quality), relevant (Relation), sufficiently informative (Quantity) and clearly conveyed (Manner), for instance by the use of an appropriate content (and thereby utterance form) or via implicature. Following most of the literature, I will not formalize the maxim of Manner, which is required primarily to constrain the relation between primary intents and the semantic contents that can serve to clearly convey them. Instead, I will adopt the common simplification that semantics more or less delivers the primary intents directly, at least on the informational side: the assumed primary informational intents for the examples in this paper will essentially be as delivered by a bare-bones, truth-conditional semantics.[2] On the attentional side there will be more to say, however; although the assumed attentional intents will be essentially compatible with existing semantic proposition-set notions in the literature (e.g., Alternative Semantics for disjunction, Alonso-Ovalle 2006), I will suggest in Section 3.3 (summarizing the more detailed presentation in Westera 2017b), inspired by work on anaphora, that also a more minimal semantic backbone can suffice, because pragmatics itself often sufficiently constrains attentional intents for them to be reliably identified by an addressee.

The maxims other than Manner will play a more central role, and will be precisely defined, broadly in line with the formal pragmatic literature. I will henceforth refer to the Gricean maxims of Quality, Quantity and Relation as the I(nformation)-maxims because they operate on informational intents,

---

2 The simplifying assumption that formal semantics would deliver the primary intents directly, without much of a role for Manner, is often left implicit, but it is apparent for instance in the common practice of applying the maxims of Quality, Relation and Quantity directly to the semantic content. But it is really the intent, that is, what is intended to be communicated, that ought to be true, relevant and sufficiently informative; semantic contents are merely an instrument to communicate intents, and for this the contents themselves need to be neither relevant nor true (sarcasm being an extreme example) — nor truth-conditional, for that matter (Bach 2006a).

anticipating the definition of a set of A(ttention)-maxims in Section 3 on attentional intents which form the main novelty of Attentional Pragmatics. Although the A-maxims will do most of the work in this paper — exhaustivity will derive from the attentional maxim A-Quantity — making both sets of maxims explicit will allow for a more precise comparison between the neo-Gricean approach and Attentional Pragmatics. To clarify, the A-maxims are not meant to replace the I-maxims; rather, a single utterance can (and, for assertions, will) express both informational and attentional intents, where each intent is subject to the relevant set of maxims.

2.1.2 Relevance and questions under discussion

Following Carlson 1983, Ginzburg 1996, Roberts 1996 (reprinted as Roberts 2012) among others, compliance with the maxims will be defined relative to a question under discussion (QUD), which I will model as a set of propositions (formally of type $\langle\langle s, t\rangle, t\rangle$, see below). This embodies the assumption that the set of all pieces of information that are broadly speaking worth sharing in a given discourse is subdivided into topically and strategically organized subsets — typically paraphrasable as questions — where any given intent needs to be relevant to only one such subset, that is, be an answer to one QUD. Note that, since an utterance can have multiple intents (e.g., what is asserted (direct intent) and conversationally implicated (indirect intent)), utterances can also relate to multiple QUDs, which I will call primary QUD, secondary QUD, etc., in line with the status of the intents that address them.

I adopt the common assumption that QUDs are, by default, closed under intersection (e.g., Schulz & van Rooij 2006, Spector 2007) — if two propositions are in the QUD, then so is their intersection (conjunction) — though only if there is no reason to believe that the person who introduced the QUD already knew that the intersection was false (after all, one should not set unachievable goals, e.g., Bratman 1987, Roberts 2012; this will matter in Section 4.2). A possible motivation for this closure assumption is that, unless there is some inhibitory relation between the two goals (e.g., that if p is made common ground we would be no longer interested in q), making their intersection common ground will be a suitable goal in its own right. Moreover, the intersection of two propositions will typically exhibit whichever features p and q share by virtue of which they were grouped in the same QUD to begin with (e.g., subject matter, relation to extra-linguistic goals). Another possible motivation is that cooperative speakers may try to ensure that their

addressee will be able to pick a unique, strongest proposition in the QUD as their answer, and closure under intersection achieves exactly that.

Other than closure under intersection by default, I assume that the range of possible QUDs is not very constrained *a priori*; which propositions are worth sharing, and which of those a speaker can reasonably decide to group together in a single QUD, is very contextual. For instance, I do not assume closure under union, because if two propositions are each useful information, that doesn't mean their (less informative) disjunction is, although depending on the context it can be. Moreover, I do not assume that QUDs are necessarily closed under negation: even if information about who was at the protest and information about who was not at the protest both happen to be worth making common ground (and this is sometimes assumed to be necessarily the case, e.g., Chierchia, Fox & Spector 2012, based on Carnap 1950), there may well be reasons for a speaker to separately address the positive and the negative side (Westera 2017c), not least because negative and positive information typically serve very different purposes (Leech 1983, Horn 1989). I will briefly return to the issue of symmetry in Section 3.4.

There is, then, a distinction between relevance in a broad sense and relevance as structured and subdivided by the QUD structure chosen by a speaker, where compliance with the maxims is defined relative to latter, narrower notion, in particular the specific QUD that the speaker chose to address.[3] The latter is also what is reflected by prosodic focus in the standard way, where, roughly, the constituents that vary among the basic answers to a QUD receive an accent (Roberts 2012, Beaver & Clark 2009, building on Rooth 1992 and before; the current paper will add nothing new regarding the relation between focus and QUDs). Although I will mostly set aside the important general issue of why a speaker might choose the types of QUDs assumed in this paper over other types of QUDs — analogous open ends exist in most if not all pragmatic theories — I will show how, in the relevant cases, we as addressees or linguists can figure out what the QUD must have been, by paying attention to preceding interrogatives and focus. Note that Attentional Pragmatics will help in this regard: the new, attentional maxims to be

---

3 By an utterance or intent "addressing" a QUD, I mean that it is specifically aimed, by the speaker, at that QUD; its compliance with the maxims is to be assessed with respect to that QUD, and meta-pragmatic markers such as prosodic focus (where available) likewise reflect that specific QUD. A given intent may well be intuitively relevant to various QUDs simultaneously, but the QUD that is specifically "addressed" in this sense occupies a pragmatically privileged position.

introduced in Section 3 will effectively impose additional constraints on the relation between an utterance and its QUD. I will return to this in Section 3.3.

2.1.3 The I(nformation)-maxims

Throughout this paper I will be using Montague's (1973) Intensional Logic as a convenient formalism, though with doxastic modality (i.e., speaker belief) instead of the usual alethic modality (i.e., logical necessity). Intensional Logic is essentially two-sorted type theory (with worlds and individuals as basic types) with a number of convenient shorthands. Since the formulas in this paper will be fairly simple (and always paraphrased by English), the following quick summary should suffice. I will use bold lowercase (p, q) to refer to a proposition, modeled as a set of worlds, that is, a function from worlds to truth values, and bold uppercase (Q) to refer to a QUD, modeled as a set of propositions or, equivalently, a function from propositions (sets of worlds) to truth values. The formula $Q(p)$ can be read as 'the proposition p is contained in the QUD Q', $^{\vee}p$ as 'proposition p is true', $^{\wedge}\varphi$ as 'the proposition that φ', $\square\varphi$ as 'the speaker believes that φ' and $\diamond\varphi$ as 'the speaker considers it possible that φ'. Where this benefits readability I will also use set-theoretical notation with the usual interpretation (these can be defined within Intensional Logic as notational shorthands; Zimmermann 1989). For a thorough introduction to Intensional Logic see Gamut 1991.

The I-maxims are defined as functions which take in an informational intent p (proposition) and a QUD Q (set of propositions) and return a truth value representing whether the maxim is complied with or not:

Definition 1. *The I-maxims:*

- I-QUALITY$(p) = \square^{\vee}p$
  'Intend to share only propositions you take to be true.'

- I-RELATION$(p, Q) = Q(p)$
  'Intend to share only propositions in the QUD.'

- I-QUANTITY$(p, Q) = \forall q \left(\left(\begin{array}{c} \text{I-QUALITY}(q)\ \wedge \\ \text{I-RELATION}(q, Q) \end{array} \right) \rightarrow (p \subseteq q) \right)$
  'Intend to share all propositions in the QUD you take to be true.'

The next subsection gives an illustration; here I will state just some general results that we can derive from this definition, for instance that a speaker

who knows what the QUD is will always know whether a given intent complies with the I-maxims or not, and that if there exists a compliant intent then it is the only one (for the given QUD). For proofs see the supplementary material (sections B and C).[4]

The definition of the I-maxims follows Grice's (1975) original, informal proposal in general outline, and also borrows from subsequent implementations. In particular, the maxim of (I-)Quantity, which is the starting point of the neo-Gricean approach to exhaustivity, is equivalent in the relevant details to the version in Harnish 1976. In general, the definition of the I-maxims is not new in any respect that matters for the issues at hand; the A-maxims in Section 3 will be more novel, and will be motivated in more detail. Nevertheless, some aspects of the current definition occasionally raise a question, and I will discuss these in the remainder of this subsection.

One aspect is that I-Quantity (as in, e.g., Harnish 1976) is so demanding that so-called *mention-some* contexts can be represented only by assuming a restriction of the QUD, say, to those propositions that are the most relevant for practical purposes, or to those propositions that come to the answerer's mind first. As an account of mention-some contexts this would be incomplete, of course, without an explanation of how an addressee may figure out the intended restriction of the QUD. But as Schulz & van Rooij (2006) note such an explanation is required regardless of how one seeks to represent the mention-some/mention-all distinction, whether pragmatically or by assuming a semantic ambiguity, hence in the rest of this paper it can be set aside.

Another aspect is that, since I-Relation does not explicitly make negations, disjunctions or intersections of answers to the QUD relevant, nor indirect, partial or merely probable answers, these would either have to be assumed to be proper elements of the QUD itself, or be accounted for by other means. For instance, one could try to explain why an indirect answer can be an appropriate discourse move by assuming that although its primary intent does not directly resolve the QUD and as such does not comply with I-Relation, the utterance has a secondary intent — a conversational implicature — that does. Alternatively one could invoke (independently motivated) QUD shifts, for instance a shift from the question of what the world is like to a question of what your epistemic state is like — the latter seems to be needed to explain why *I'm not sure* is a felicitous response to basically any question.

---

4 Supplementary materials can be found at https://osf.io/dj4qm.

As for disjunctive answers, the strictness of I-Relation predicts that either the disjunction must be considered worth making common ground in its own right and as such be included in the QUD, or a disjunctive answer will not be able to comply with the maxims relative to the QUD, thus requiring hedging (*All I know is that…*) or prosodic marking (e.g., (fall-)rising intonation as a cue of uncertain relevance, indicating 'is this partial answer useful?'; Ward & Hirschberg 1985). Regarding the latter, note that being rational/cooperative does not entail full compliance with the maxims, but merely trying one's best to comply, and making sure (lest the addressee draw false inferences) that deviations from full compliance are either contextually obvious or indicated by the speaker by means of explicit hedges or prosodic cues (Grice's (1975) *opting out*).[5] The possibility of opting out and hedges will occasionally resurface in this paper, but will not matter for the main contributions.

Lastly, a number of authors have proposed probabilistic extensions or implementations of (more or less) Gricean pragmatics to do justice to, for instance, uncertainty about the world, about the QUD or about the meanings of words (Franke 2011, Russell 2012, Frank & Goodman 2012, Degen & Tanenhaus 2015, Franke & Jäger 2016, Bergen, Levy & Goodman 2016). This important strand of work is especially necessary for utterances that occur in an underspecified context, without a clear-cut QUD and perhaps lacking the kinds of intonational cues that in part reflect the structure of the QUD. In the present paper such circumstances will be mostly avoided (e.g., by specifying the prosody and by preceding the examples with an explicit question), and no probabilistic approach will be needed here, although it will ultimately be necessary for applying the theory to more naturalistic data. None of these approaches change the outlook of the standard recipe with regard to the main puzzles covered in the present paper.

---

5 This means that conversational implicatures, which rely for their communication on a presumption of compliance with the maxims, are effectively indicated by (the absence of) conventional signals. Grice's (1975) original definition of conversational implicature permits this: although it states that such implicatures rely on a presumption of compliance with the maxims ("or at least the cooperative principle"), it does not restrict the range of contextual and linguistic evidence that can in principle affect such a presumption, and indeed Grice himself recognizes the possibility of both contextual and explicit cancellation. Note that this does not nullify the distinction between conversational implicatures and conventional implicatures, the main defining feature of which is, rather, that the former are indirect intents (relying on the recognition of another speech act and assumptions about its compliance with the maxims) while the latter are direct intents (i.e., communicated directly by conventional means, not piggy-backing on another speech act).

2.2 The neo-Gricean approach to exhaustivity

Exhaustivity has been modeled pragmatically in what has come to be known as the *neo-Gricean approach*. To illustrate, consider again the example from the introduction, now with an explicit question preceding it:

(2)   Who (of Alf, Beth, Cee) was at the protest?
         ALF or BETH was at the protest. *(Implied: not both, not Cee)*

The neo-Gricean approach is based on the following intuition: if the speaker had believed the stronger information that both Alf and Beth attended, then according to the maxim of I-Quantity they would have asserted that — since the speaker did not, they must believe it is false (and similarly for combinations with Cee).

   To see how this works formally, we need to specify the primary informational intent p for (2), that is, what it asserts, and its QUD Q (where superscript $\cap, \cup$ on the QUD indicate closure under intersection and union):

$$p = {}^{\wedge}(Pa \vee Pb) \qquad Q = \{{}^{\wedge}Pa, {}^{\wedge}Pb, {}^{\wedge}Pc, \ldots\}^{\cap, \cup}$$

Before proceeding, let us consider briefly how we as linguists, or as addressees, can figure out that this is the right QUD for the utterance (and also how much we might still figure out without the preceding interrogative). In the scope of this paper I often rely on an explicit interrogative and/or focus intonation for determining the QUD, or at least for constraining it sufficiently to enable the computation of clear, falsifiable predictions. With regard to example (2), note that we can determine (already without the preceding interrogative) that the QUD must contain the propositions denoted by the disjuncts and their intersection and union. The QUD must contain the disjunction as a whole, because otherwise its informational intent (i.e., ${}^{\wedge}(Pa \vee Pb)$) could not comply with I-Relation. This disjunction could be included in the QUD for its own sake, or as a partial answer composed of two more basic propositions in the QUD, namely the two individual disjuncts. Prosody strongly suggests the latter: the intended prosody for (2), with focus on the names, conveys that propositions in the QUD vary in exactly those places, and a QUD containing propositions of the form 'X was at the protest' is the minimal way to achieve this.[6] The intersection of these disjuncts is added in accordance

---

6 But not the only way: the prosody is also compatible with a disjunctive multi-wh QUD like 'Who or who attended the protest?'. But although this type of QUD could perhaps occur in the right sort of context (e.g., as an echo question in response to a disjunctive assertion in which you misheard the two names), it is odd out of the blue, so here I will set it aside.

with the default assumption of closure under intersection, and their union is included because otherwise the intent p would not have complied with I-Relation relative to the QUD.

These constraints on the QUD suffice of course for the theory to predict exhaustivity, for instance, that not both were at the protest, and that, *if* someone else's presence was relevant too, they must have been absent. The explicit interrogative, in particular its explicit domain comprising of Alf, Beth and Cee, suggests that the latter holds for Cee — even without making any assumptions about the attentional intent or QUD of the interrogative. In fact, we could have avoided the reliance on an interrogative entirely: we can replace the interrogative by an explicit disjunction and observe the same exhaustivity effects:

(3)   Alf was at the protest, or Beth, or Cee, or some combination, is all I know.
       (Yes, in fact) ALF or BETH was at the protest.

For this reason I will not say much about the semantics or attentional intents of wh-interrogatives in this paper; it is not needed for the theory to make falsifiable predictions. (In fact, we could use the correspondence in exhaustivity effects between (2) and (3) to figure out what the attentional intent of wh-interrogatives ought to be like.)

Anyway, with the assumed QUD and informational intent we can prove that the following holds, in a certain relevant subclass of models (namely, those models which validate the definitions of the maxims, along with a number of principles such as the introspection axioms for belief and that the speaker knows what the QUD is):

$$\text{I-QUANTITY}(p, Q) \rightarrow \neg \Box Pab$$

where Pab abbreviates $Pa \wedge Pb$. That is, if the utterance complies with I-Quantity, the speaker must not have the belief that both Alf and Beth were at the protest.

As is well known, this I-Quantity implication is almost exhaustivity, but not quite: from the absence of a positive belief ($\neg \Box Pab$) one may not in general conclude the presence of a negative belief ($\Box \neg Pab$) — the speaker could simply be ignorant, after all. That the I-Quantity implication falls short of exhaustivity was already pointed out by Soames (1982: p. 534), in a discussion of Gazdar's (1979) account. More recently, Sauerland (2004) called the gap between I-Quantity and exhaustivity, that is, between not-believing and

believing-that-not, the *epistemic step*. To explain how participants in a conversation take the epistemic step, Soames proposed that they normally assume each other's *opinionatedness* (Horn (2001) notes that this proposal can be found already in Mill 1865). Indeed, if we assume for whatever reason that the speaker in (2) takes herself to know whether or not both were at the protest, the I-Quantity implication can be strengthened, yielding exhaustivity:

$$\neg\Box Pab \ \wedge \ \overbrace{(\Box Pab \vee \Box\neg Pab)}^{\text{opinionatedness assumption}} \ \rightarrow \ \Box\neg Pab$$

I-Quantity plus opinionatedness is the essence of the neo-Gricean approach (e.g., Horn 1972, Gazdar 1979, Sauerland 2004, Schulz & van Rooij 2006, Spector 2007, Geurts 2011). As I mentioned, in Section 4 I will discuss three empirical challenges for the neo-Gricean approach: its inability to predict exhaustivity in two cases where the maxim of I-Quantity isn't operative, namely questions and a quizmaster's hints, and contexts where the required opinionatedness assumption cannot be granted.

To clarify, the neo-Gricean approach, or any of the approaches discussed in this paper for that matter, applies not just to disjunctions but to any case where the QUD contains a stronger proposition than what is mentioned in the utterance, including so-called scalar cases, such as *some* implying 'not all' and *warm* implying 'not hot'. I focus on disjunction in this paper only for the sake of exposition; disjunction makes it easier, for instance, to control for the QUD in the relevant respects.

2.3 The grammatical approach to exhaustivity

The neo-Gricean approach has been and continues to be adopted by many, but the last decade or two have witnessed the rise of a contender, the *grammatical approach*. The grammatical approach assumes covert operators in the grammatical structure of a sentence that directly render their syntactic argument exhaustive relative to some grammatically constrained set of alternatives (e.g., Fox 2007, Chierchia, Fox & Spector 2009, 2012, Katzir & Singh 2014, Mayr & Romoli 2016). For instance, for disjunctions such as (2) the grammatical approach can assume an exhaustivity operator on each disjunct:

(4)   EXH(Alf was at the protest) or EXH(Beth was at the protest).

The result of interpreting this structure is roughly paraphrasable as *Only Alf was at the protest or only Beth*, which indeed has the desired 'not both' implication. Exhaustivity operators and grammatically constrained alternatives have turned out to afford the grammatical approach the flexibility to take on an impressive range of data.

It is important to acknowledge that the grammatical approach still relies on pragmatics in various ways. For one, exhaustivity is still typically assumed to involve the exclusion of *relevant* alternatives, relying on pragmatics for a notion of relevance — even though a grammatical mechanism can be assumed to further restrict or even modify the pragmatic alternatives (e.g., Katzir 2007).

Second, although grammatical exhaustivity operators suffice to add exhaustivity to the main semantic content, this does not let us draw any inferences unless this content is perceived to serve the communication of an informational intent that complies with the maxim of I-Quality: only then can we conclude that the speaker believes that the exhaustivity is true (and, if we assume the speaker's beliefs are accurate, that it is indeed true). I will return to this facet in Section 4.2 when considering the puzzle of exhaustivity in questions, which lack an informational intent and are not subject to I-Quality, after all.

Third, the grammatical approach relies on pragmatics for governing, at least in part, the placement of exhaustivity operators. For the sake of concreteness, let me mention two such principles (not all implementations rely on the same principles, or define them in the same way). One is the Strongest Meaning Hypothesis (based on Dalrymple et al. 1998), that in the face of multiple possible readings (e.g., with and without an exhaustivity operator) we should prefer the stronger one (Chierchia, Fox & Spector 2012). I will return to the Strongest Meaning Hypothesis in Section 4.3, where I investigate its relation to the opinionatedness assumption in the neo-Gricean approach. Another principle is Hurford's Constraint (after Hurford 1974), that disjuncts should not be informationally redundant. Hurford's Constraint can force the insertion of exhaustivity operators in a disjunction (Singh 2008, Chierchia, Fox & Spector 2012, Katzir & Singh 2014, Fox & Spector 2018). For instance:

(5)   Alf was there, or both Alf and Beth.

Here the second disjunct entails the first and hence does not contribute to the overall informational content/intent (logically, $p \vee (p \wedge q) \equiv p$), *unless*

we interpret the first disjunct exhaustively as meaning 'only Alf' by applying an exhaustivity operator to it.

In Section 4, I will compare the neo-Gricean and grammatical approaches to Attentional Pragmatics on the three empirical puzzles introduced above: exhaustivity on the hints of a quizmaster, exhaustivity on questions, and exhaustivity without an opinionatedness assumption. Although this particular selection of puzzles was made with a comparison between the two pragmatic approaches as the main goal, actually the first and the third (in part) arose in the literature on the grammatical approach, as arguments against the neo-Gricean approach and thereby, lacking a pragmatic alternative, in support of a grammatical approach. Hence, the current paper can be framed as agreeing with the grammatical approach that these puzzles falsify the neo-Gricean approach, but, by presenting a pragmatic alternative that does handle them, undermining their use as positive arguments for a grammatical approach.

This possible framing notwithstanding, the current paper will not address the main argument that has been presented in favor of the grammatical approach, namely, its ability to account for cases where exhaustivity effects seem to exist in a grammatically embedded position. For instance:

(6)   Cee believes that ALF or BETH was at the protest.
                    (Implied: Cee believes that not both were there)

The utterance has a reading where the exhaustivity effect 'not both' is itself embedded under belief, that is, 'Cee believes that not both were there'. This would be problematic for the neo-Gricean approach, and arguably pragmatic approaches more generally, since pragmatics operates on intents — communicative intentions — and these are (by and large) carried by entire utterances and not by their parts (but cf. Simons 2011). The reason why the present paper will not deal with embedded exhaustivity is that Attentional Pragmatics and the neo-Gricean approach, both being fundamentally pragmatic approaches, offer essentially the same outlook in this regard. Whatever I (or anyone) might have to say about Attentional Pragmatics in relation to embedded exhaustivity will not depend on the particulars of Attentional Pragmatics, hence it would be better left for a separate contribution on pragmatics vs. grammar in general.[7] Indeed, the particular selection of puzzles in this paper was made with a comparison between the two pragmatic ap-

---

7 For the sake of concreteness: there are various possible pathways to embedded exhaustivity, besides grammatical exhaustivity operators, whose potential has hardly been explored (Geurts 2011), e.g., different types of pragmatic implicature, pragmatic impliciture (Bach 1994; also 'explicature', Recanati 2001), typicality inference (e.g., van Tiel 2014), meta-

proaches as the main goal, although I will also consider how the grammatical approach fares in each case.

3 Attentional Pragmatics

3.1 Definition of the A(ttention)-maxims

Parallelling the informational notions, let an *attentional intent* be a set of things to which the speaker intended to draw attention (in the sense of a communicative intention), and let an *attentional content* be the set of things to which an utterance draws attention by virtue of its semantics. In these terms, that utterances have attentional contents is uncontroversial; it is the assumption that utterances also have attentional intents that sets Attentional Pragmatics apart from most work (an exception being Ciardelli, Groenendijk & Roelofsen 2009). Attentional intents being communicative intentions, they should be governed by a set of constraints in their own right, namely a set of attentional maxims or A-maxims, defined below.

To clarify, the A-maxims are assumed to operate alongside the I-maxims, not to replace them. As we will see in Section 4, the addition of the A-maxims solves precisely the three puzzles this paper sets out to solve, in a way that the I-maxims alone cannot: they predict exhaustivity on the hints of a quizmaster (exempt from I-Quantity but, as I will argue, not A-Quantity), on questions (which, I propose, lack an informational intent but do have an attentional intent), and without relying on a contextual opinionatedness assumption (because the maxim of A-Quantity will be more demanding than I-Quantity, it will license sufficiently strong inferences as it is).

Attentional Pragmatics will be concerned only with attention drawn to *propositions*, which are the kinds of things that can be relevant, true, possible, etc. Formally, let an attentional intent be a set of propositions, the same type of object as a QUD ($\langle\langle s,t\rangle,t\rangle$). A minimal assumption about an attentional

---

linguisticness, lexical semantics, and interactions of these mechanisms. For instance, different embedding verbs give rise to different kinds of exhaustivity effects, suggesting that the lexical semantics of these predicates has at least some role to play (e.g., Uegaki 2015), either instead of or in interplay with grammar and/or pragmatics. It is an open and in part empirical issue whether the various embedded and unembedded exhaustivity effects can and should be treated as a unified phenomenon at all (e.g., Russell 2006, Geurts 2011, Simons 2011). The grammatical approach suggests that they (by and large) should be, while a pragmatic approach to unembedded exhaustivity will likely require a more pluriform treatment of embedded cases.

intent would be to assume that it must be some subset of all propositions to which the utterance draws attention. Which subset this is could be fixed by an appropriate semantics (cf. the discussion in Section 2.1), but in the next subsection I argue why this may not be necessary. There I will also motivate the particular attentional intents assumed for the examples in this paper.

The A-maxims follow a similar recipe as the I-maxims introduced above, relative to a QUD Q, this time for an attentional intent A:

Definition 2. *The A-maxims:*

- A-QUALITY$(A) = \forall a(A(a) \rightarrow \Diamond(^{\vee}a \wedge \forall b((b \subset a \wedge A(b)) \rightarrow \neg^{\vee}b)))$
  'Intend to draw attention only to propositions that you consider possible, in particular, possible independently of anything stronger to which you intend to draw attention.'

- A-RELATION$(A, Q) = \forall a(A(a) \rightarrow Q(a))$
  'Intend to draw attention only to relevant propositions.'

- A-QUANTITY$(A, Q) = \forall a\left(\left(\begin{array}{c} \text{A-QUALITY}(\{a\}) \wedge \\ \text{A-RELATION}(\{a\}, Q) \end{array}\right) \rightarrow A(a)\right)$
  'Intend to draw attention to all relevant propositions you consider independently possible.'

As on the informational side, some general results can be proven, for instance that a speaker who knows what the QUD is will always know whether their attentional intent complies with the A-maxims, and that if there is a compliant attentional intent for a given speaker, then it is the only one (for the given QUD, and given that speaker's beliefs). For proofs see the supplementary material (sections B and C). Note that, if all elements of the attentional intent are logically independent (no entailment relation), then A-Quality can be simplified to just the part before the conjunction: $\forall a(A(a) \rightarrow \Diamond^{\vee}a)$.

As conversational maxims, the A-maxims should be understandable as aspects of general rationality. To see that this is so, let us conceive of the relevant propositions (i.e., those in Q) as pertinent goals to pursue, namely, pieces of information to be made common ground in the current stretch of discourse. It will generally be beneficial for a rational agent to keep track of all and only pertinent goals that are potentially achievable and that might need to be independently pursued (i.e., that will not come for free by pursuing some other goal). This is, in essence, what A-Quality, A-Relation and

A-Quantity jointly require, and it clearly applies not just to discourse goals but also to, say, the goals involved in cooking a meal or building a house.

Moreover, similar constraints have been formulated in the prior literature. The first part of A-Quality (up to the ∧, i.e., that propositions in the attentional intent be considered possible) is assumed by Roelofsen (2013), building on Ciardelli, Groenendijk & Roelofsen 2009, who calls it "Attentive Sincerity" (also "Interactive Sincerity" in Coppock & Brochhagen 2013); it corresponds also to "Genuineness" in Zimmermann 2000: p. 270, and to the "Viability" constraint in Biezma & Rawlins 2012 p. 46; a similar intuition underlies Gazdar's (1979) "clausal implicatures". Biezma & Rawlins moreover assume a constraint like A-Relation, though incorporated as a presupposition in their question semantics; Simons (2001) assumes a comparable "relatedness condition" on disjunctions, and in fact the same idea is found already in Grice 1989, who states that disjunction serves to specify possibilities "that relate in the same way to a given topic". Given A-Quality and A-Relation, A-Quantity is simply necessary to complete the definition: while A-Quality and A-Relation state only what *not* to draw attention to, and as such can be vacuously complied with by an empty attentional intent, A-Quantity sets a lower bound. Lastly, there is a partial correspondence between A-Quantity and Dayal's (1996) influential "strongest true answer" principle for questions, according to which the proposition set denoted by a question should always contain a proposition that is true and entails all the other true propositions in the set.[8]

The second part of A-Quality can be considered a weaker version of the descriptive generalization about disjunctions known as Hurford's Constraint (Hurford 1974), according to which disjuncts must not entail each other (cf. "independence" in Zimmermann 2000, p. 720). A-Quality is weaker in permitting one disjunct to entail the other, provided the weaker disjunct is at least considered possible independently of the stronger one. Note that unlike A-Quality, Hurford's Constraint is (at first sight) falsified by felicitous disjunctions such as *Alf, Beth, or both*, among many other examples (see Potts 2013: p. 24).[9] This independence requirement of A-Quality can be motivated in terms of attentional economy: conceiving of the relevant proposi-

---

8 More precisely, Dayal's principle follows from A-Quantity for the attentional intents expressed by declaratives, and in part for the attentional intents expressed by interrogatives (namely those interrogatives which make their QUD 'fully explicit', in some sense) but space limitations preclude demonstrating this here.

9 Recall from Section 2.3 that proponents of the grammatical approach often do assume Hurford's Constraint, and regard the felicity of certain examples as evidence for the local, ex-

tions again as representing pertinent goals, it is simply unnecessary to keep track of a goal a in addition to a more specific goal b, if you know that a will automatically be achieved, and can only be achieved, together with b. This aspect of A-Quality matters only for so-called Hurford disjunctions such as (9), discussed further below, including in particular those like (9a) with "distant entailing disjuncts" (Fox & Spector 2018).

3.2 Basic predictions with respect to exhaustivity

Let us consider how exhaustivity is predicted for the simple disjunction from the introduction, repeated here:

(7)   Who (of Alf, Beth, Cee) was at the protest?
        ALF or BETH was at the protest.   *(Implied: not both)*

To apply Attentional Pragmatics we need a primary attentional intent A, in addition to the informational intent p and QUD Q given earlier:[10]

$$Q = \{^\wedge Pa, ^\wedge Pb, ^\wedge Pc\}^{\cap,\cup} \qquad p = {}^\wedge(Pa \vee Pb) \qquad A = \{^\wedge Pa, ^\wedge Pb\}$$

That is, I assume that the speaker intends to draw attention to the propositions expressed by the two disjuncts — for justification of this assumption see Westera 2017b, along with a more general characterization of attentional intents in terms of sentence form.

With the given QUD and attentional intent for (7) the following can be proven:

$$\text{A-QUANTITY}(A, Q) \rightarrow \Box \neg Pab$$

Concise proof: that no attention is drawn to Alf and Beth's joint presence ($^\wedge Pab \notin A$) implies, by contraposition of the conditional of A-Quantity, that it must either not be relevant ($\neg\text{A-RELATION}(\{^\wedge Pab\}, Q)$) or not be considered possible independently of anything stronger in the attentional intent ($\neg\text{A-QUALITY}(\{^\wedge Pab\})$), and since we know it is in fact relevant, and that there is nothing stronger in A, the explanation can only be that the speaker must not consider it possible. That is, assuming compliance with the maxim of A-Quantity entails that the speaker must believe that not both Alf and Beth were at the protest, which is of course the exhaustivity implicature we set out to explain. In a similar way we can derive Cee's absence ($\Box \neg Pc$).

---

haustive interpretation of the weaker disjunct, which would break the entailment relation between the disjuncts (e.g., Chierchia, Fox & Spector 2012, Katzir & Singh 2014).

10 For brevity, let superscript $\cap, \cup$ on the QUD indicate closure under intersection and union.

Attentional Pragmatics furthermore predicts that one way to prevent an exhaustivity implicature is by including a third disjunct *...or both* in order to draw attention to Alf and Beth's joint presence:

(8)   Alf was at the protest, or Beth, or both.

Here we must assume the same QUD and primary informational intent as before, but a richer attentional intent (due to the additional disjunct):

$$Q = \{^\wedge Pa, ^\wedge Pb, ^\wedge Pc\}^{\cap, \cup} \qquad p = {^\wedge}(Pa \vee Pb) \qquad A = \{^\wedge Pa, ^\wedge Pb, ^\wedge Pab\}$$

By drawing attention to Alf and Beth's joint presence, this utterance satisfies the consequent of A-Quantity for this proposition and even implies, through A-Quality, that the speaker considers their joint presence possible, which is the polar opposite of exhaustivity ($\Diamond Pab \equiv \neg\Box\neg Pab$). Altogether, the theory does justice to the motivating intuition given in the introduction: that exhaustivity arises when a relevant proposition is not mentioned. (Besides this, recall that the traditional I-maxims are still assumed to be active too, still allowing for the usual pragmatic implications, e.g., that the speaker believes that the disjunction is true.)

To illustrate A-Quality, and especially the independence requirement it contains, consider a context where the speaker knows that either Alf is at the protest or all three of Alf, Beth and Cee, but never only two of them. For this speaker, uttering (9a) is fine while uttering (9b) is strange:

(9)   *The speaker knows if Alf and Beth were at the protest, so was Cee.*

    a.   Alf is at the protest, or Alf, Beth and Cee.

    b. ?? Alf is at the protest, or Alf and Beth, or Alf, Beth and Cee.

Example (9b) is bad precisely because the speaker does not consider the second disjunct possible independently of the third, and including it falsely gives the impression that they do. The part of A-Quality after the conjunction captures exactly this. Formally, assuming the following QUD and primary intents for (9b):

$$Q = \{^\wedge Pa, ^\wedge Pb, ^\wedge Pc\}^{\cap, \cup} \qquad p = {^\wedge}Pa \qquad A = \{^\wedge Pa, ^\wedge Pab, ^\wedge Pabc\}$$

The following can be proven for (9b):

$$\text{A-Quality}(A, Q) \rightarrow \Diamond(Pab \wedge \neg Pabc)$$

This is the inference that is incompatible with the intended context in (9). Meanwhile, for the variant in (9a) without the middle disjunct we can derive exactly the opposite, from A-Quantity, assuming the same QUD but an attentional intent without the middle disjunct ($A = \{^\wedge Pa, ^\wedge Pabc\}$):

$$\text{A-QUANTITY}(A, Q) \rightarrow \neg \Diamond (Pab \wedge \neg Pabc)$$

This is a type of exhaustivity inference, derived from A-Quantity. What is noteworthy here is that the proposition $^\wedge Pab$ is only partially excluded, namely, only the part which it does not share with $^\wedge Pabc$. Katzir & Singh (2014) argue, for an example isomorphic to this (their (8)), that the partial exclusion of a proposition like this is particularly challenging for pragmatic accounts of exhaustivity, and indeed it is for the neo-Gricean approach (i.e., based on the I-maxims alone), which is insensitive to the individual disjuncts of an utterance.

For a more general specification of the predictions of Attentional Pragmatics an exhaustivity operator can be defined as a shorthand for what compliance with A-Quantity implies (Westera 2017a). It can be defined as follows:

$$\text{EXH}(A, Q) = \bigcap_{\substack{a \in Q \\ a \notin A}} \left(\overline{a} \cup \bigcup_{\substack{b \in A \\ b \subset a}} b \right)$$

See the supplementary material (Section D) for a proof that this definition is a shorthand for the implications derived from A-Quantity; see Westera 2017a for its application to additional examples. In certain circumstances, this operator is formally conservative with respect to some existing operators, in particular the operator in Schulz & van Rooij 2006 — for a formal statement and proofs of these results see again the supplementary material (Section E). Schulz & van Rooij's operator was stipulated to take the individual disjuncts into account, but only for the sake of descriptive coverage, without a pragmatic motivation. In fact, they present their operator as a refinement of a more basic operator which they derive from the neo-Gricean approach (I-Quantity plus opinionatedness), and it is precisely the refinement step — making their operator sensitive to the individual disjuncts — that is not pragmatically motivated. Attentional Pragmatics can be seen as retroactively showing how their operator can be derived from pragmatic principles after all, but, crucially, from A-Quantity, without relying on either opinionatedness or I-Quantity.

3.3 Remarks about the assumed QUDs and intents

I will briefly address (but, for reasons of space, mostly defer to other works) two important, seemingly open ends of the current theory: how to identify the QUD for a given utterance, and how to identify its purported attentional intent. Clearly, merely stipulating QUDs and intents on a case-by-case basis is not sufficient for a falsifiable theory.

First, regarding the identification of QUDs: why a speaker might choose one kind of QUD (e.g., the types of QUDs assumed in this paper) over others is an important open question for most QUD-based accounts, as already mentioned in Section 2.1.2. But setting aside the why, figuring out what the QUD must have been like, for a given utterance, turns out to be fairly straightforward for the examples in the current paper, provided focus prosody is taken into account and domain restrictions are made explicit, and further facilitated by preceding the relevant examples with explicit questions. This was illustrated in Section 2.2, justifying the assumed QUD for our basic type of example, that is, a disjunction with focus on the disjuncts. Now, the theory presented here does not claim to enable the identification of QUDs *in general*, so this remains an important research direction; it means that the predictions of Attentional Pragmatics on naturally occurring examples will inevitably be less precise. But the theory at least presents a step forward in this regard, because the A-maxims effectively impose an additional set of constraints on the relation between utterances and their QUDs.

A case where the latter is particularly clear is the following, related to the well-known "symmetry problem" that arises when a QUD contains some propositions as well as their negations:

(10)   Q: Who was and who wasn't at the protest (of Alf, Beth, Chris, ...)?
        A: (Of those people,) ALF was there. [final fall]

Relative to the symmetrical QUD arguably introduced by the explicit question, A's response, which asserts (and draws attention to) only a single, positive proposition (Alf's presence), may well comply with the I-maxims (namely if they are ignorant about Beth and Chris), but cannot comply with the A-maxims. This is because, according to A-Quantity, since the speaker must consider either Beth's presence or Beth's absence possible (or both), and likewise for Chris, drawing attention to neither of these possibilities violates A-Quantity — if, that is, the response indeed addresses the explicitly introduced, symmetrical QUD. In other words, Attentional Pragmatics predicts for (10) that either speaker A violated a maxim (in which case we might ex-

pect hedges or special prosodic marking, such as a final rise), or they must have chosen to address only the positive, asymmetrical sub-QUD explicitly.[11] For more detailed discussion of these predictions I must refer to Westera 2017c. What I wanted to highlight in the current section is that Attentional Pragmatics imposes additional constraints on QUDs, restricting the range of possible analyses for a given example further than the I-maxims alone.

As for the second seemingly open issue, namely the identification of attentional intents, we will likewise see that Attentional Pragmatics already constrains the notion more than it might at first appear. On the informational side, I mentioned in Section 2.1 the common simplification that truth-conditional semantics would directly deliver the primary informational intents, essentially bypassing the Manner maxim by conflating intent with content. We could adopt a similar simplification on the attentional side and have attentional intents be delivered directly by a suitable semantics. Indeed, the attentional intents assumed in this paper are broadly in line with existing formal semantic approaches that assign proposition sets to sentences, such as Hamblin alternatives for disjunction (Alonso-Ovalle 2008) or unrestricted inquisitive semantics (Ciardelli, Groenendijk & Roelofsen 2009). The reader is free to adopt such a framework for easy computation of conditional predictions, that is, conditioned on the assumption that the semantic framework of choice indeed delivers the correct attentional intents. However, in Westera 2017b: ch. 6 I argue that this is unnecessary, at least for the range of examples considered in the present paper, and list a number of advantages of such a semantically non-committal approach (though without making a definitive argument against adopting a proposition-set semantics, with which Attentional Pragmatics is in principle perfectly compatible). That is, Attentional Pragmatics itself, even without a designated semantic backbone for attentional intents, often sufficiently constrains the range of possible attentional intents for the theory to yield clear, falsifiable predictions.

3.4 Summary

Attentional Pragmatics is based on the assumption that drawing attention to relevant propositions is a communicative intention in its own right, alongside assertion, governed by its own set of maxims, the A-maxims. We have

---

11 This also aligns with the focus marking in (10): if the QUD addressed by A had contained negated propositions too, one would have expected an additional, 'verum'-type focus on the verb).

already seen how this addition to basic Gricean pragmatics (i.e., the I-maxims alone) enables new analyses, such as for Hurford disjunctions like (9), and with regard to maxim violations in the case of a symmetrical QUD (10). The next section describes in detail how the A-maxims also enable novel pragmatic solutions to the three problems introduced in Section 1: exhaustivity on the hints of a quizmaster (in a nutshell: it can be argued that these are exempt from I-Quantity but not A-Quantity), exhaustivity on questions (again, subject not the I-maxims but to the A-maxims), and exhaustivity without an opinionatedness assumption (because A-Quantity, unlike I-Quantity, is strong enough to deliver exhaustivity without it).

4 Comparison with regard to three puzzles

4.1 Exhaustivity on the hints of a quizmaster

4.1.1 The empirical picture

The first of the three challenges for the neo-Gricean approach is the following puzzle, from Fox 2014. When a quizmaster is giving a hint to a contestant, the quizmaster is not expected to share all the information they have — hints are partial answers by definition. Another way of framing this is that hints are exempt from the maxim of I-Quantity: not all true relevant information needs to be shared. And yet, surprisingly for the neo-Gricean approach which relies on this maxim, exhaustivity can still be present:[12]

(11)   Quizmaster: (Of these three boxes over here,) there is money in box A
        or in box C.   *(implied: not in both, and not in B)*

Exhaustivity is present here: if the quizmaster had known that there was money in box B, or in both A and C, they could be accused of misleading the contestant.[13] And yet no genuine I-Quantity implication is present here to

---

12 Example (11) should be read, as before, with falling intonation and focus on the two disjuncts. In particular, it is important to rule out an alternative intonation that seems to me very natural for hints: fall-rise (Constant 2012, Wagner 2012, Westera 2019), which is a signal of partial answerhood and (therefore) non-exhaustivity.

13 Perhaps the 'not both' implication is more robust in the following variant, as mentioning another intersection ('A and D') could make closure of the QUD under intersection more likely, and hence the relevance of 'A and C':

(i)   Quizmaster: There is money in box A, or in boxes A and D, or in box C.
                                    (More robustly predicted: not in both A and C)

the effect that, for instance, the quizmaster necessarily lacks the belief that the money is in box A — in fact the quizmaster may well know exactly that that's where the money is. Accordingly, as Fox notes, there is no way for the neo-Gricean approach, which relies crucially on I-Quantity, to predict exhaustivity for (11). Moreover, the reasonable assumption that the quizmaster is at least *pretending* to be complying with I-Quantity cannot help in this case, as Fox notes: if the I-Quantity implication is pretended, there is no way for an implication that is derived from it to not also be merely pretended — but the exhaustivity is genuine.

4.1.2 Explanation in Attentional Pragmatics

Since Attentional Pragmatics derives exhaustivity from A-Quantity instead of I-Quantity, one could try to account for (11) by supposing that, while quizmasters are not supposed to provide all the information they possess, perhaps they do have to mention all possible choices for a quiz participant (i.e., comply with A-Quantity), lest they be found guilty of misleading. If true, this would deliver exhaustivity as before, through A-Quantity. However, while I find it quite intuitive that a quizmaster's hints would be exempt from I-Quantity but not from A-Quantity, this is not entirely satisfactory unless we can truly explain the different exemption status of the different maxims.[14] Puzzlingly, it does not seem to instantiate a broader pattern of I-maxims vs. A-maxims, because the maxims of I/A-*Quality* seem to behave in the opposite way: while (11) is exempt from A-Quality — the quizmaster need not actually consider it possible that there is money in box A — it does need to comply with I-Quality, that is, there has to money in box A or box C.

As an explanation for this exemption/non-exemption pattern, and for what is going on in (11), I propose the following:

- **Assumption:** Quizmasters may pretend to know *less* about the quiz answers than they actually do, but not to know *more* about the quiz answers than they actually do (i.e., information that may be false).

This seems to me a safe assumption to make; one could imagine reading this as part of the job instructions for a quizmaster. Crucially, it follows that pragmatic implications of not-knowing ($\neg\square$) need not be genuine, whereas

---

14 Moreover, the intuition that a quizmaster isn't exempt from A-Quantity is not really to be trusted: it may well be a mere consequence of the explanandum, i.e., that hints imply exhaustivity, rather than an explanation for it.

pragmatic implications of knowing ($\square$) must be genuine, and this accounts precisely for the apparent exemption status of the different maxims:

- I-Quality: implies knowing ($\square$), hence hints are subject to it.

- I-Quantity: implies not-knowing ($\neg\square$), hence hints are exempt.

- A-Quality: implies not-knowing ($\diamond$, which is equivalent to $\neg\square\neg$), hence hints are exempt.

- A-Quantity: implies knowing ($\neg\diamond$, which is equivalent to $\square\neg$), hence hints are subject to it.

Given Attentional Pragmatics, the prediction that hints are exempt from I-Quantity whilst subject to A-Quantity solves precisely the puzzle posed by (11), where exhaustivity is present without a genuine I-Quantity implication. In fact, the foregoing line of explanation can work for any theory that delivers the knowing-that-not implication of exhaustivity directly, rather than via a not-knowing implication as in the neo-Gricean approach. This is true of Attentional Pragmatics, but potentially also of the grammatical approach, discussed next.

4.1.3 Hints in the grammatical approach

The crucial part of the foregoing explanation is the assumption that a quiz-master's implied knowledge about the quiz answers must be genuine, and this will include exhaustivity regardless of whether it is delivered by A-Quantity or by a grammatical operator plus I-Quality. However, the exhaustivity operator and I-Quality represent only part of the story for the grammatical approach: its predictions depend also on whether the principles and contextual assumptions that guide the placement of exhaustivity operators apply to a quizmaster. This is currently an open question, not addressed by Fox (2014), and it is difficult to answer, as the principles used in the grammatical approach vary across implementations. For the sake of concreteness I will discuss a more traditional principle and a more recent approach.

An influential principle has been the Strongest Meaning Hypothesis (following Dalrymple et al. 1998), which states that exhaustivity operator placements which strengthen the overall meaning are preferred (e.g., Chierchia,

Fox & Spector 2012, Mayr & Romoli 2016).[15] We could try to rely on this principle for favoring the exhaustive reading in (11). However, for this to actually solve the quizmaster puzzle we should motivate why this principle, and not I-Quantity, can be safely assumed to apply to a quizmaster. It is not obvious that such motivation can be found: if we know that the quizmaster will pretend to know less than they actually do, should we really be 'charitable listeners' and read as much as possible into their hints? The answer to this may depend on whether we conceive of the Strongest Meaning Hypothesis as a pragmatic constraint which may vary across contexts, or as a hard architectural or conventionalized constraint in the grammar itself.

A less standard implementation of the grammatical approach, Meyer's (2013) *Matrix K* theory, seems to avoid this issue. It assumes that every matrix sentence is in fact the complement of an implicit doxastic operator K ('the speaker believes that…'). This offers various compositions of exhaustivity operators — below K, above K, or both — resulting in different readings. According to this approach, the quizmaster example (11) has distinct readings entailing either ignorance or exhaustivity, expanding the grammatical approach's general pattern of replacing cooperativity-based pragmatic enrichment by contextually constrained disambiguation. Meyer argues that disambiguation towards the exhaustive reading can take place simply by virtue of our contextual assumption/knowledge that the quizmaster cannot in fact be ignorant. The hybrid grammatical-pragmatic approach in Buccola & Haida 2019, which likewise derives both the ignorance reading and the exhaustive reading in the grammar, would enable a similar account. I will return to this strand of grammatical approaches in Section 4.3, in relation to the opinionatedness assumption. For now, note that my assumption that a quizmaster's hints are exempt from certain maxims and not others has no real bearing on the grammatical approach: if exhaustivity is a grammatically generated semantic entailment, its truth will be implied via I-Quality, to which I argued a quizmaster's hints *are* subject.

Wrapping up, the foregoing corroborates Fox's arguments that cases like (11) are a problem for the neo-Gricean approach, a problem which can *in principle* be resolved by the grammatical approach — the latter depending on the exact nature of the principles that guide the placement of exhaustivity op-

---

15 The Strongest Meaning Hypothesis is not universally adopted in the grammatical approach (e.g., see Chemla & Spector 2011, fn.6 for an empirical argument against; see also Poortman 2017), and sometimes it is formulated differently; see Section 4.3 for a more detailed discussion.

erators, and/or on whether the traditional I-Quantity inferences are instead derived in the grammar. More central to the aims of this paper, I have argued that Fox's puzzle is not a problem for pragmatic approaches per se: if we assume that quizmasters are allowed only to pretend to know less than they actually do, not more, then Attentional Pragmatics generates the right predictions.

4.2 Exhaustivity in disjunctive questions

4.2.1 The empirical picture

Questions have been noted to imply exhaustivity (e.g., Bartels 1999, Roelofsen & van Gool 2010, Geurts 2011, Biezma & Rawlins 2012):[16]

(12)   Was ALF or BETH at the protest?   *(implied: not both)*

The neo-Gricean approach cannot account for this, because questions are exempt from I-Quantity, indeed, from all the I-maxims: questions do not have a primary informational intent to which these maxims could apply, or in other words they lack assertive force (e.g., Frege 1918). By contrast, Attentional Pragmatics arguably does generalize to questions: although questions do not serve to provide information, they arguably do serve to draw attention to things, and as such have attentional intents subject to the A-maxims just like assertions. Thus, since (12) is subject to A-Quantity just like the assertion in (2) with which this paper started, one might expect it to imply exhaustivity in exactly the same way: since no attention is drawn to Alf and Beth's joint presence, the speaker must not consider it possible. This is not completely correct, as I will explain further below, but it already shows the potential for Attentional Pragmatics to offer a more or less uniform account of exhaustivity in questions and assertions.

This is not to say that there are no differences between exhaustivity on questions and exhaustivity on assertions (and in fact a more uniform approach along the foregoing lines can explain such differences, as we will see). For instance, it is often assumed that in the case of assertions the 'not both' implication, and exhaustivity more generally, is not just an implication but a conversational implicature, that is, an indirectly communicated informational intent, hence part of what the speaker meant (or at-issue). By contrast,

---

16 As before, the disjunction needs to be read with focus on the disjuncts and a final fall, i.e., the type of intonation triggering what some call an 'alternative question' interpretation.

the same implication in the case of questions is commonly assumed to be a presupposition (Bartels 1999, Aloni & Égré 2010, Biezma & Rawlins 2012): it is implied by the utterance to be true but not part of what the speaker meant, that is, not an informational intent, and non-at-issue. The contrast between assertions and questions regarding the status of exhaustivity is widely assumed, across different strands in the literature, so I will just take it for granted in what follows (though for an illustration of a possible empirical consequence of this contrast see Westera 2017b: ch. 12). A theory of exhaustivity on assertions and questions should of course account for this contrast, but ideally it would explain the obvious parallelisms too, for instance how exhaustivity is affected, on questions and assertions alike, by disjuncts like *or both*, by focus, and by rising intonation.

Existing accounts of both the 'not both' implication of disjunctive questions and its non-at-issue status have described rather than explained it, and have done so independently of theories of exhaustivity on assertions. For instance, Biezma & Rawlins (2012) encode an exhaustivity presupposition into the core semantics of interrogatives, without attempting to derive it from a pragmatics of questions, and also without trying to find a common basis for exhaustivity on questions and assertions. Similarly for the influential approach of Dayal (1996), which is based on the assumption that the proposition set denoted by a question should always contain a strongest true answer. Skipping the details, this principle results in a 'not both' presupposition in the case of disjunctive questions, but also, for instance, a uniqueness presupposition in the case of singular wh-questions (e.g., *which boy was sleeping?* presupposes there is only one such boy), among other applications related to constraints on question formation (e.g., Fox & Hackl 2006, Abrusán & Spector 2010) — it has been empirically fruitful, but its explanatory potential, say, its status as a pragmatic principle, is unclear (see also footnote 8).

4.2.2 Account in Attentional Pragmatics

I assume the following about questions (by which I mean, in this context, utterances of interrogatives):

- Questions and assertions alike have primary attentional intents, subject to the same A-maxims.

- Questions are different from assertions in lacking a primary informational intent (although they may in principle have secondary informational intents, e.g., implicatures).

- Interrogative syntax is essentially a marker indicating the absence of a primary informational intent. Besides that, interrogatives have the same semantics as their declarative counterparts, and also focus reflects the QUD they address in the same way.

So basically, questions are like assertions minus the assertion speech act, leaving only the attentional component.[17] This idea has not, to my awareness, been considered in the literature, but it is a natural option in Attentional Pragmatics: once we assume that questions have attentional intents, they do not need anything else in order to serve the pragmatic function they do. After all, by drawing attention to relevant propositions one would effectively be indicating the desire that they be made common ground, which is conceivably the primary purpose of (information-seeking) questions — see Biezma & Rawlins 2012 for an account of questions in line with this view (and the semantic contents they assume are equivalent to the required attentional intents in the present paper). In fact, this view on questions is not limited in scope to the exhaustivity effects of disjunctive questions; see Westera 2017b: ch. 12) for an account along these lines of the differences between so-called 'polar questions' and 'alternative questions', as well as other combinations of focus structure and final contour. In any case, with this view on questions in place, and assuming the same type of QUD and attentional intent as for the corresponding assertions, Attentional Pragmatics would predict exhaustivity for (12) in exactly the same way as for the corresponding assertion, through A-Quantity.

The foregoing account of the disjunctive question in (12) is not quite right, however. For it to work, the QUD must contain Alf and Beth's joint presence ($^\wedge Pab$), which in Section 2.1.2 I motivated, following previous work, by assuming that QUDs are typically closed under intersection due to the structure of relevance and/or the cooperativeness of ensuring a unique answer possibility for the addressee. However, closure of QUDs under intersection is a mere

---

17 The same need not be assumed for so-called embedded questions, as in *John wonders whether Alf or Beth was at the protest*, which arguably semantically denote questions rather than pragmatically ask them. Following Roberts (2012) and Biezma & Rawlins (2012), I do not assume the Equivalence Thesis of Belnap 1982 according to which the semantic content of unembedded questions would be equivalent to that of the analogous embedded questions.

default assumption, assumed to be true of QUDs in the absence of reasons to believe otherwise. And precisely such a reason is present in (12): the speaker who utters the interrogative disjunction is the one introducing the QUD, and they would not have included a given proposition in their QUD if they already knew it to be false.[18] Hence, the exhaustive interpretation of (12), that not both Alf and Beth were there, does not quite follow from A-Quantity; only that their joint presence could not reasonably have been in the QUD (for if it had been included, *then* A-Quantity would have implied that the speaker considered it false, hence that it should not have been included — a contradiction).

Let me go through this explanation with a bit more precision, starting from the attentional intent of (12):[19]

$$\textbf{Attentional intent:} \{^\wedge Pa, ^\wedge Pb\}$$

Suppose, counterfactually (to derive a contradiction), that Alf and Beth's joint presence ($^\wedge Pab$) is an element of the QUD in (12):

$$\textbf{QUD (not really):} \{^\wedge Pa, ^\wedge Pb, ^\wedge Pab\}$$

In that case, since no attention was drawn to their joint presence, compliance with A-Quantity would imply that the speaker must consider it false — this would be exhaustivity in the usual manner. But if the speaker already knew the intersection was false, then — since they are the one introducing the QUD — why would they have included it? Therefore, contrary to the initial supposition, Alf and Beth's joint presence cannot be in the QUD in (12), that is, the QUD must be like this:

$$\textbf{QUD:} \{^\wedge Pa, ^\wedge Pb\}$$

And this entails that exhaustivity must be explained slightly differently: the speaker must not consider Alf and Beth's joint presence possible, not because it is in the QUD and no attention was drawn to it, but because it *isn't* in the QUD, and they *would have* included it in the QUD (and drawn attention

---

18 I motivated this assumption in Section 2.1.2, i.e., that a QUD should not contain propositions that the person who introduces the QUD already knows are false. It aligns with Roberts 2012, and more generally with the idea that one should not set goals that one already knows are unachievable (found for instance in Bratman 1987).

19 For motivation for the assumed attentional intent on the basis of certain semantic assumptions, see Westera 2017b: ch. 6.

to it) had they considered it possible, given default closure under possible intersections.[20]

The foregoing suggests that whereas exhaustivity in the case of assertions is the exclusion of relevant propositions, exhaustivity in the case of questions is the exclusion of propositions that would have been relevant had they been deemed possible. Both explanations rely crucially on A-Quantity and the fact that no attention was drawn to Alf and Beth's joint presence, but they diverge from there, based on the fact that closure under intersection applies differently when the speaker is the one introducing the QUD. The idea that exhaustivity in questions is primarily about unmentioned propositions not being relevant (rather than, as for assertions, unmentioned relevant propositions not being true) is independently motivated in Biezma & Rawlins 2012, and the resulting overall explanation for the 'not both' implication is similar to the one sketched in Groenendijk & Roelofsen 2009, namely, that if the speaker had considered 'both' relevant they would have listed it as a possible answer. But if the end result is the same for assertions and questions, that is, a 'not both' implication, then what difference does it make?

The contrast in at-issueness (presupposition vs. implicature) As it turns out, the foregoing account predicts that 'not both' is at-issue in the case of assertions but not questions, as follows (a summary of Westera 2020b). In the case of assertions, 'both' is relevant, that is, establishing it is a conversational goal. For this reason, establishing its negation 'not both' would effectively serve to prune an unachievable goal and thereby keep the conversation focused on those which are achievable.[21] The information 'not both' therefore serves a meta-communicative purpose, or what I will call a *secondary goal*, in line with the notion of "secondary content" (intent, in the present sense) of Gutzmann & Turgay 2019, which covers the meta-communicative kinds of contributions of, for instance, discourse particles, conventional implicatures

---

20 One might wonder if leaving the decision of whether to include the 'both' proposition in the QUD entirely to the speaker (as opposed to enforcing closure under intersection as a general constraint) will not lead to misunderstandings. What if the addressee thinks that 'both' *is* in the QUD, given that QUDs are by default closed under intersection? But this situation will not occur: if it had been in the QUD, then the speaker should have drawn attention to it (or used final rising intonation or a more explicit hedge to indicate that there are other relevant possibilities besides those mentioned) — which they did not.

21 This view on the conversational purpose of the 'not both' implication aligns with Horn 1989: the typical reason for conveying negative information is the prior consideration of its positive counterpart — we tend not to be interested in negative information in its own right.

and intonational meaning. Thus, the type of goal served by the information 'not both' is secondary in the sense that keeping the set of discourse goals tidy is not the primary aim of the discourse.

Now, I assume that the organization of discourse goals into QUDs applies across the board, to both primary and secondary discourse goals (put differently: any type of informational intent has to address an appropriate QUD, whether it is asserted or implicated). It then makes sense for all secondary goals related to pruning unachievable goals from the primary QUD to be grouped in single, topically coherent secondary QUD. This is the type of QUD which the 'not both' implicature could serve to address, and it explains why assertions have this secondary informational intent. By contrast, in the case of questions 'both' is not contained in the primary QUD to begin with, hence there is no reason to prune this non-existing goal, and accordingly there is no secondary QUD containing 'not both'. Therefore, in the case of questions, there is no QUD for the 'not both' implication to compliantly address, and therefore it cannot serve the communication of an intent; it is merely an implication/inference.

Summing up, 'not both' is part of what is meant (in particular, a secondary intent) in the case of assertions because 'both' was primarily relevant and therefore 'not both' was secondarily relevant for the sake of keeping the primary goal set tidy, whereas in the case of questions 'both' was not primarily relevant to begin with (for reasons identified above), hence 'not both' will not be relevant, either. This difference notwithstanding, let me repeat that ultimately the exhaustivity of both questions and assertions derives from A-Quantity. Indeed, the main point of this section was that questions are subject to A-Quantity but not I-Quantity, which lets Attentional Pragmatics extend to questions where the neo-Gricean approach does not.

4.2.3   Question exhaustivity in the grammatical approach

Questions have only recently begun to draw some attention in the grammatical approach. For instance, Ciardelli & Roelofsen (2017) suggest that Hurford's Constraint, one of the core principles in the grammatical approach, holds for both questions and assertions. However, they do not seek to account for the exhaustivity implications of questions, and indeed I am unaware of approaches that do (extensive work on exhaustivity in *embedded* so-called questions notwithstanding, e.g., Gajewski & Sharvit 2012).

The grammatical approach treats exhaustivity as part of the informational (semantic) content, and as such directly predicts 'not both' to be part of what is meant in the case of assertions, namely, as part of what is asserted (the primary informational intent), and its truth is implied accordingly by the maxim of I-Quality.[22] But in the case of questions nothing is asserted, so the maxim of I-Quality doesn't have anything to apply to (or alternatively, in some accounts interrogativity turns the informational content/intent into a tautology), and accordingly any exhaustivity generated below the interrogativity does not end up being implied to be true. A more technical way of putting this is that entailments do not 'project' out of interrogativity, so neither does exhaustivity as conceived in the grammatical approach. It is technically possible of course to define a special-purpose exhaustivity operator for interrogatives, one which makes exhaustivity escape interrogativity as a presupposition (Biezma & Rawlins (2012) define such an operator, though independently of the grammatical approach), but such special treatment is not very explanatory.

Summing up, Attentional Pragmatics accounts for exhaustivity in a way that generalizes from assertions to questions, while also predicting some differences, in particular that the 'not both' implication serves to communicate an intent in the case of assertions but not questions. Meanwhile, the neo-Gricean approach does not generalize to exhaustivity in questions, and it is not clear at present either how the grammatical approach would handle it.

---

22 The fact that the grammatical approach treats exhaustivity as part of the primary informational intent as opposed to an implicature marks a major deviation from prior characterizations of exhaustivity, but it has not received much attention, perhaps in part because it has been obfuscated by the continued use of the term "implicature" in the grammatical approach (for a discussion of terminological confusions surrounding implicature see Bach 2006b, Geurts 2011). Major as this deviation may be, it has not to my awareness been empirically assessed or otherwise defended other than by showing that making exhaustivity part of the informational content allows one to uniformly capture *the overall truth-conditions* (disregarding the assertion/implicature distinction) of various unembedded and embedded exhaustivity. This is a great result, but one should wonder whether the distinction between main and secondary intent (implicature) can in the end somehow be recovered from the grammatical approach and, if not, what (if anything) we would loose.

4.3 Exhaustivity without an opinionatedness assumption

4.3.1 The empirical picture

Recall that the neo-Gricean approach relies on an opinionatedness assumption (e.g., for (2), the assumption that $\Box Pab \vee \Box \neg Pab$), because the I-Quantity implication on its own ($\neg \Box Pab$) falls short of exhaustivity ($\Box \neg Pab$) — this is the epistemic step (Sauerland 2004). By contrast, Attentional Pragmatics predicts exhaustivity without relying on an opinionatedness assumption; the A-Quantity implication is strong enough as it is. The reason for this is that A-Quantity is more demanding than I-Quantity: it requires that attention be drawn not just to propositions that are believed to be true, but to any proposition that is considered even merely possible. One might wonder whether this is an advantage or disadvantage: what is the explanatory value of an opinionatedness assumption to begin with, and, empirically speaking, can exhaustivity occur without it?[23]

The explanatory value of the opinionatedness assumption is unclear at best. To my awareness it has not been argued to be an independently reasonable assumption, and the neo-Gricean approach's reliance on an opinionatedness assumption has been criticized by some for being *ad hoc* (Groenendijk & Stokhof 1984, Chierchia, Fox & Spector 2012, Westera 2014, Chierchia 2017). Indeed, it simply does not seem rational to go about the world assuming, without evidence, that our interlocutors will be opinionated about whatever the QUD happens to be. Nevertheless, the opinionatedness assumption has been embraced by many (e.g., Soames 1982, Horn 1984, Matsumoto 1995, Green 1996, Russell 2006, Schulz & van Rooij 2006, Geurts 2011; see Scontras & Tessler 2018 for a probabilistic implementation of opinionatedness in the RSA framework).

This appreciation for the epistemic step is due largely to the fact that exhaustivity implications seem to disappear if the speaker denies their own opinionatedness, as Soames (1982) noted:

---

23 A similar assumption plays a role in Zimmermann's (2000) account of modal disjunctions (e.g., *You may do A or B*). The interpretation of such disjunctions is affected by a contextual *authority* assumption, namely, that the speaker is in a position to set permissions and obligations for the addressee. My arguments in this section are independent of this, and, more generally, although my proposal will deny that an opinionatedness assumption is necessary for exhaustivity, it does not deny that contextual assumptions of various kinds can interfere with pragmatic inferences. Indeed, an (un)opinionatedness assumption is still expected to affect exhaustivity in my account (as I will explain shortly), and the interpretation of modal disjunctions may well be affected by an authority assumption.

(13) A: Who (of Alf, Beth, …) was at the protest?

  B: I'm not sure about Alf, but Beth was there.

Speaker B's hedge, denying their opinionatedness about Alf, causes the usual exhaustivity implication 'not Alf' to be absent. Similarly explicit examples can be found throughout the literature. Breheny, Ferguson & Katsos (2013) rely not on explicit cues but on context for manipulating the speaker's (un)opinionatedness, and present experimental results that seem to support the reliance of exhaustivity implications on an opinionatedness assumption. They presented participants with visual scenes and descriptions whose exhaustive interpretation was expected to affect eye movements around the scene. The visual scenes also showed a speaker, whose opinionatedness was manipulated by visibly occluding the relevant items from the speaker. When the participants knew that the speaker had only partial knowledge about the scene, the exhaustivity implications were significantly weaker than when the participants did not know this. Goodman & Stuhlmüller (2013) conduct similar experiments, showing comparable results.

  However, there is reason to believe that these experimental results as well as Soames's original observation have been misinterpreted. After all, in the relevant examples (such as (13), but also the more contextualized cases of Breheny, Ferguson & Katsos 2013 and Goodman & Stuhlmüller 2013), not only is the supposed opinionatedness assumption removed; it is in fact *replaced by an un-opinionatedness assumption*, or even knowledge to that effect. And since to be *un*opinionated is to be unable to give an exhaustive answer, it is unsurprising that knowing in advance that the speaker is *un*opinionated has some effect on the perceived exhaustivity implications, regardless of whether their presence would have depended on an opinionatedness assumption to begin with. After all, when we know that the speaker is not opinionated about Alf, we know in advance that we must not interpret the answer *Beth* as excluding Alf's presence, and we will try to find a way of interpreting the utterance that is compatible with this uncertainty. For instance, we may interpret the main part of B's response in (13) as addressing an implicitly more restricted QUD: *but of the people I do know about, Beth was there*. As long as there is a way to accommodate the speaker's unopinionatedness, there will be a reading of the utterance that lacks exhaustivity — and this is true even if exhaustivity itself does not depend on an opinionatedness assumption.

  A better test case for the neo-Gricean approach's reliance on an opinionatedness assumption would be to deny not the speaker's opinionatedness, but merely the opinionatedness *assumption*, and see if exhaustivity still appears.

The following example is such a test case (I considered a similar example in Westera 2014):

(14)   A: I *may* be asking the wrong person — you probably don't know this — but do you have any idea who was at the picnic?

        B: John was there, Bob, Mary, and Sue.   *(with falling intonation)*

        A: So you do know! Great, thanks. Only four though, that's a bit disappointing…

Here, A explicitly refrains from assuming B's opinionatedness, though crucially (unlike (13)) without assuming their *un*opinionatedness instead, and in the end A still takes B's answer to be exhaustive. When presenting an analogous Dutch example (but without A's final response) to nine informants, in a simple questionnaire with auditory stimulus, eight agreed that B's response implies exhaustivity (with strength 5 out of 5; Westera 2017b). See Dieuleveut, Chemla & Spector 2019 for a more proper experiment supporting the same empirical conclusion: exhaustivity does not rely on an opinionatedness assumption. They conclude by saying that this finding mandates either a grammatical approach, or a pragmatic approach that (by involving higher-order beliefs, as in the Rational Speech-Act framework) lets speakers act *as if* opinionatedness is assumed even if they know it isn't. Leaving aside the question of how plausible such pretense would really be, Attentional Pragmatics can crucially do without. Either way, the foregoing results outweigh the earlier misinterpreted evidence supposedly in favor of an opinionatedness assumption, especially combined with it having been *ad hoc* from the start.

4.3.2   Is opinionatedness conveyed by B's response?

Another way to frame what is going on in (14) is that B's response itself *conveys* opinionatedness, rather than presupposing a contextual assumption to that effect. Because exhaustivity entails opinionatedness, this is formally true regardless of which comes first in the derivation, as it were. But more than a formal truth, I think there is also a strong intuition regarding (14), that what lets us infer B's opinionatedness is the fact that B is answering so utterly straightforwardly; B simply sounds like they know exactly what they are talking about. What to make of this intuition in Attentional Pragmatics, where exhaustivity does not rely on opinionatedness? And could we not try to amend the neo-Gricean approach by building on this intuition, that opinionatedness is conveyed rather than assumed? I will argue that Attentional

Pragmatics in fact explains precisely this intuition, and that amending the neo-Gricean approach in this regard is not straightforward and, when properly pursued, will likely lead to something close to Attentional Pragmatics.

First, consider how Attentional Pragmatics explains the above intuition. If B had not been sure about the others not being at the picnic, then according to Attentional Pragmatics they should have either (i) drawn attention to the other possibilities, or (ii) explicitly restricted the QUD ("of those people which I know about, …"), or (iii) signaled a possible A-Quantity violation ('opting out') with some kind of hedge or, in the theory of Westera 2018, final rising intonation. Hence, according to Attentional Pragmatics, B's response implies exhaustivity, hence also opinionatedness, indeed by virtue of it being such a straightforward, confident response.

Crucially, unlike the neo-Gricean approach, Attentional Pragmatics explains this without having to rely on opinionatedness as some kind of default assumption, the explanatory value of which I have pointed out is dubious. Instead, opinionatedness is shown to follow from A-Quantity, which is a constraint that is not a mere reformulation of opinionatedness, but which can be understood as a trait of rational behavior more generally: it is important to keep track of the propositions that are to be made (and can still be made) common ground, and this is important and rational regardless of how opinionated one is.

Moreover, A-Quantity delivers 'the right amount' of opinionatedness: as Schulz & van Rooij (2006) note, the exact notion of opinionatedness involved in exhaustivity should be made sensitive somehow to the things mentioned in the utterance. For instance, for the plain disjunction in (2) the speaker is opinionated about Alf and Beth's joint presence but not about Alf's individual presence, or Beth's for that matter — in fact the speaker must be *un*opinionated about those. Thus, to say that B in (14) simply sounds opinionated is not enough to amend the neo-Gricean approach; we need a more principled account of the pragmatic import of the 'things mentioned in an utterance' — and Attentional Pragmatics provides such an account: the speaker sounds exactly as opinionated as compliance of their attentional intent with A-Quantity entails they are.[24]

This is achieved, moreover, without *ad hoc* assumptions about intonational meaning. Some authors have suggested that opinionatedness (i.e., exactly the right amount of it, subject to what is being mentioned) is conveyed

---

24 Just to reiterate, this does not prevent an *un*opinionatedness assumption from still affecting exhaustivity, as in (13).

directly by means of final falling intonation (Hara 2005, Schulz & van Rooij 2006); or that ignorance (again, exactly the right amount of it) is conveyed by a certain different intonation (Buccola & Haida 2019). But this idea has not been worked out in the context of a broader theory of intonational meaning and pragmatics, making its explanatory potential difficult to assess. By contrast, a long strand of work on intonational meaning characterizes rising and falling intonation as indicating not the speaker's opinionatedness but whether the utterance is pragmatically a "complete" contribution to the discourse (e.g., Pierrehumbert & Hirschberg 1990), a notion which in Westera 2018 is made more precise as indicating full compliance with the maxims. By introducing the maxim of A-Quantity into this picture, Attentional Pragmatics directly explains how the final fall, as a signal of compliance with the maxims, would convey the right amount of opinionatedness, with no new assumptions needed about intonational meaning.

4.3.3   Opinionatedness in the grammatical approach

Although proponents of the grammatical approach (e.g., Chierchia, Fox & Spector 2012, Chierchia 2017) have recognized the opinionatedness assumption as a challenge for the neo-Gricean approach, there has been comparatively little reflection on whether the grammatical approach itself is able to truly avoid it. Superficially this may seem obvious: the grammatical approach bypasses I-Quantity and opinionatedness altogether by generating exhaustivity directly by a grammatical operator. But, as in the previous subsection, it really depends on which principles one assumes for guiding the placement of exhaustivity operators to begin with.

Let us consider the Strongest Meaning Hypothesis again, which is the supposed preference for stronger meanings in the face of multiple possible readings (though see below for slightly different formulations). Even on the surface, taking the speaker to *mean* as much as possible (i.e., as much as is compatible with the uttered sentence) is not very different from taking the speaker to *believe* as much as possible, where the latter is a rephrasing of opinionatedness (following Schulz & van Rooij 2006). Indeed, the Strongest Meaning Hypothesis follows, informally at least, from opinionatedness plus I-Quantity: the maxim of I-Quantity requires the speaker to mean as much as possible given their beliefs, and given opinionatedness the speaker must believe as much as possible given their utterance, therefore the speaker must

have meant as much as possible given their utterance (Strongest Meaning Hypothesis).

Moreover, the Strongest Meaning Hypothesis seems just as *ad hoc* as the opinionatedness assumption: just as it would not be rational to go about the world assuming (without evidence) that everyone is maximally opinionated, it does not seem rational to just opt for the strongest possible interpretation of a given observation (linguistic or otherwise). From a probabilistic perspective, one should opt not for the strongest interpretation but for the most likely one (indeed, see Poortman 2017 for evidence against the Strongest Meaning Hypothesis and in favor of a 'Maximal Typicality Hypothesis', concerning the interpretation of reciprocal constructions).

The Strongest Meaning Hypothesis is sometimes defined not as a preference for the insertion of exhaustivity operators that strengthen the meaning, but as a prohibition on the insertion of exhaustivity operators that would *weaken* it (or do nothing at all). For instance, Fox & Spector (2018) incorporate the latter type of constraint in their 'Economy Condition'. The mere prohibition of weakening operators (as opposed to a preference for strengthening operators) is not in general enough to decide in favor of the stronger, exhaustive reading (i.e., to derive that the exhaustivity operator must be there). Accordingly, such approaches need to rely on additional principles, contextual assumptions (such as opinionatedness in the neo-Gricean approach), or the absence of hedges or intonational cues (as I discussed above). As another example of the latter, Chierchia (2017) writes about the source of opinionatedness in the grammatical approach that a speaker can convey that an exhaustified parse is intended by means of a 'confidence display', though without defining the latter term or explaining how it would display exactly the right amount of confidence given the things mentioned in the utterance (as discussed above).

Now, the Strongest Meaning Hypothesis (however defined) is not universally adopted in the grammatical approach, and there are of course other principles or mechanisms that may guide the insertion of exhaustivity operators. As mentioned in Section 4.1, a non-standard implementation of the grammatical approach is Matrix K theory by Meyer (2013), which seeks to derive not just exhaustivity itself but also inferences pertaining to speaker beliefs directly in the grammar. This is pursued further by Buccola & Haida (2019), who assume that relevance is closed under negation and belief (if p is relevant, so is $\neg p$, the speaker's belief that p, their belief that $\neg p$, etc.). They show that, relative to such a set of alternatives, Gricean (I-)Quantity-based

reasoning often leads to the only possible reading for a given utterance being an exhaustive one, without the need for an opinionatedness assumption. To nevertheless explain cases where exhaustivity is absent, they informally point to a certain special intonation, in line with approaches based on an opinionatedness assumption discussed above in Section 4.3.2 (and subject to the exact same criticism).

In their paper, Buccola & Haida frame their assumptions of closing relevance under negation and speaker belief as "small adjustments" to prior views on relevance, but this is not quite true from the present perspective. While it may be plausible that if p is relevant in a broad sense then so is $\neg p$ or 'believe p', it is quite another thing to assume that these are therefore relevant to the same degree or in the same way, let alone that they enter into pragmatic inference in the same way. Even though inferences about a speaker's beliefs obviously matter for successful communication (hence, e.g., I-Quality), this does not mean that a discourse is itself *about* those beliefs, and that our inferences about them are necessarily part of what the speaker *meant*. A core assumption of the QUD framework, as explained in Section 2.1.2, is that speakers present their utterances as addressing a certain subset of all broadly relevant propositions, namely the QUD. The QUD is linguistically marked for instance by focus, in a way that does not seem compatible with assuming that QUDs are generally closed under negation and belief — even if relevance in a broad sense might be. This linguistic marking makes QUD-based inferences more robust, and it lets us isolate and formalize one part of pragmatics — the relation between an utterance and its QUD(s) — while leaving the remainder informal and partial for now (such as why a speaker might choose one QUD rather than another) without such incompleteness rendering our theories unfalsifiable (see Section 3.3). Buccola & Haida depart from this view not in their assumptions about closure of relevance in a broad sense per se, which may be valid (indeed, I assume the same in Westera 2017b), but in their assumption that exhaustivity crucially *depends* on this broader notion of relevance — and that is not a "small adjustment" at all. According to Attentional Pragmatics, exhaustivity depends only on the QUD.[25]

---

25 Buccola & Haida (2019) also rely, additionally, on a more formal set of alternatives, namely grammatically constrained alternatives to which the exhaustivity operators in the grammar would be sensitive (hence it is only the more pragmatic component of their approach that relies on relevance in a broad sense). But also this more constrained set of alternatives does not quite align with the notion of QUD, formally nor conceptually, not least because, by

Most of the foregoing does not argue against (or in favor of) Buccola & Haida's (2019) approach; and the QUD-based approach itself is not without important open issues. What it does show is that the approaches are couched in such different frameworks, with no direct mapping between their various moving parts, that a more detailed comparison has to be left to future work. Pending that, the main conclusion from this section is that, although proponents of the grammatical approach have (rightly) criticized the neo-Gricean approach for its reliance on an *ad hoc* and empirically inadequate opinionatedness assumption, an alternative pragmatic approach exists that does not need it. Moreover, a generally accepted solution within the grammatical approach itself seems to be lacking — although the account of Buccola & Haida 2019 does reveal a possible approach.

5 Conclusion

This paper presented Attentional Pragmatics, a theory based on the idea that exhaustivity arises when relevant propositions are not mentioned, or more precisely, when the speaker did not intend to draw attention to them. By contrast, the neo-Gricean approach is based on the more common idea that exhaustivity would arise when relevant propositions are not *asserted*. This at first sight minor difference results in very different and, I argued, better predictions for a range of challenges for the neo-Gricean approach, of which I could discuss three in the scope of this paper: exhaustivity on the hints of a quizmaster, exhaustivity on questions, and exhaustivity without an opinionatedness assumption. This suggests that, if exhaustivity is a pragmatic phenomenon, it is a matter of attention rather than of information, that is, it derives from the maxim of A-Quantity, not I-Quantity.

Two of the three challenges for the neo-Gricean approach discussed in this paper turned out not to be definitively resolved in the grammatical approach either, namely exhaustivity on questions and the opinionatedness assumption, inviting future work in this regard. On the other hand, two topics that are central to the grammatical approach could not be addressed in detail in this paper at all. One is Hurford disjunctions, which are disjunctions where one disjunct entails the other. Although several examples in this paper were in fact Hurford disjunctions, these served only to explain how the attentional maxims work. For a detailed comparison between Attentional

additionally assuming belief operators in the grammar (in line with the Matrix K theory of Meyer 2013), these end up in some of the grammatical alternative sets as well.

Pragmatics and the neo-Gricean and grammatical approaches on this subject I refer to Westera 2017a, 2020a, where the latter covers a more representative sample of Hurford disjunctions (unembedded and embedded). The second topic that was not covered in any detail is embedded exhaustivity, though for a different reason: Attentional Pragmatics simply does not change the pragmatic outlook on embedded exhaustivity compared to the neo-Gricean approach, which is that it will require a more pluriform treatment than what the grammatical approach offers.

References

Abrusán, Márta & Benjamin Spector. 2010. A semantics for degree questions based on intervals: Negative islands and their obviation. *Journal of Semantics* 28(1). 107–147. https://doi.org/10.1093/jos/ffq013.

Aloni, Maria & Paul Égré. 2010. Alternative questions and knowledge attributions. *The Philosophical Quarterly* 60(238). 1–27. http://www.jstor.org/stable/40468275.

Alonso-Ovalle, Luis. 2006. *Disjunction in alternative semantics.* Amherst, MA: University of Massachusetts dissertation. https://www.proquest.com/docview/305305643.

Alonso-Ovalle, Luis. 2008. Innocent exclusion in an alternative semantics. *Natural Language Semantics* 16(2). 115–128. https://doi.org/10.1007/s11050-008-9027-1.

Bach, Kent. 1994. Conversational impliciture. *Mind and Language* 9(2). 124–62. https://doi.org/10.1111/j.1468-0017.1994.tb00220.x.

Bach, Kent. 2006a. The excluded middle: Semantic minimalism without minimal propositions. *Philosophy and Phenomenological Research* 73(2). 435–442. http://www.jstor.org/stable/40040979.

Bach, Kent. 2006b. The top 10 misconceptions about implicature. In Betty J. Birner & Gregory Ward (eds.), *Drawing the boundaries of meaning: Neogricean studies in pragmatics and semantics in honor of Laurence R. Horn* (Studies in Language Companion Series 80), 21–30. John Benjamins Publishing Company. https://doi.org/10.1075/slcs.80.03bac.

Bach, Kent & Robert Harnish. 1979. *Linguistic communication and speech acts.* Cambridge, MA & London: MIT Press.

Bartels, Christine. 1999. *The intonation of English statements and questions: A compositional interpretation.* New York, NY: Routledge.

Beaver, David & Brady Clark. 2009. *Sense and sensitivity: How focus determines meaning* (Explorations in Semantics 12). Hoboken, NJ: John Wiley & Sons. https://doi.org/10.1002/9781444304176.

Belnap, Nuel D., Jr. 1982. Questions and answers in Montague Grammar. In Esa Saarinen & Stanley Peters (eds.), *Processes, beliefs, and questions* (Studies in Linguistics and Philosophy 16), 165–198. Springer. https://doi.org/10.1007/978-94-015-7668-0_6.

Bergen, Leon, Roger Levy & Noah D. Goodman. 2016. Pragmatic reasoning through semantic inference. *Semantics and Pragmatics* 9(20). https://doi.org/10.3765/sp.9.20.

Biezma, María & Kyle Rawlins. 2012. Responding to alternative and polar questions. *Linguistics and Philosophy* 35(5). 361–406. https://doi.org/10.1007/s10988-012-9123-z.

Bratman, Michael E. 1987. *Intentions, plans, and practical reason*. Cambridge, MA: Harvard University Press.

Breheny, Richard, Heather J. Ferguson & Napoleon Katsos. 2013. Taking the epistemic step: Towards a model of on-line access to conversational implicatures. *Cognition* 126(3). 423–440. https://doi.org/10.1016/j.cognition.2012.11.012.

Buccola, Brian & Andreas Haida. 2019. Obligatory irrelevance and the computation of ignorance inferences. *Journal of Semantics* 36(4). 583–616. https://doi.org/10.1093/jos/ffz013.

Carlson, Lauri. 1983. *Dialogue games: An approach to discourse analysis* (Studies in Linguistics and Philosophy 17). Dordrecht: Reidel. https://doi.org/10.1007/978-94-015-3963-0.

Carnap, Rudolf. 1950. *Logical foundations of probability*. Chicago, IL: University of Chicago Press.

Chemla, Emmanuel & Benjamin Spector. 2011. Experimental evidence for embedded scalar implicatures. *Journal of Semantics* 28(3). 359–400. https://doi.org/10.1093/jos/ffq023.

Chierchia, Gennaro. 2017. Scalar implicatures and their interface with grammar. *Annual Review of Linguistics* 3. 245–264. https://doi.org/10.1146/annurev-linguistics-011516-033846.

Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2009. Hurford's constraint and the theory of scalar implicatures: Evidence for embedded implicatures. In Paul Egré & Giorgio Magri (eds.), *Presuppositions and implicatures: Proceedings of the MIT-Paris workshop* (MIT working papers in linguistics 60), 47–62. Cambridge, MA: MIT Press.

Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2012. Scalar implicature as a grammatical phenomenon. In Claudia Maienborn, Paul Portner & Klaus von Heusinger (eds.), *Semantics: An international handbook of natural language meaning*, vol. 2 (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 33), 2297–2331. Mouton de Gruyter. https://doi.org/10.1515/9783110253382.2297.

Ciardelli, Ivano, Jeroen Groenendijk & Floris Roelofsen. 2009. Attention! *Might* in inquisitive semantics. *Semantics and Linguistic Theory (SALT) 19*. https://doi.org/10.3765/salt.v19i0.2520.

Ciardelli, Ivano & Floris Roelofsen. 2017. Hurford's constraint, the semantics of disjunction, and the nature of alternatives. *Natural Language Semantics* 25(3). 199–222. https://doi.org/10.1007/s11050-017-9134-y.

Constant, Noah. 2012. English rise-fall-rise: A study in the semantics and pragmatics of intonation. *Linguistics and Philosophy* 35(5). 407–442. https://doi.org/10.1007/s10988-012-9121-1.

Coppock, Elizabeth & Thomas Brochhagen. 2013. Raising and resolving issues with scalar modifiers. *Semantics and Pragmatics* 6(3). 1–57. https://doi.org/10.3765/sp.6.3.

Dalrymple, Mary, Makoto Kanazawa, Yookyung Kim, Sam Mchombo & Stanley Peters. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* 21(2). 159–210. https://doi.org/10.1023/A:1005330227480.

Dayal, Veneeta. 1996. *Locality in WH quantification: Questions and relative clauses in Hindi* (Studies in Linguistics and Philosophy 62). Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-011-4808-5.

Degen, Judith & Michael K. Tanenhaus. 2015. Processing scalar implicature: A constraint-based approach. *Cognitive science* 39(4). 667–710. https://doi.org/10.1111/cogs.12171.

Dieuleveut, Anouk, Emmanuel Chemla & Benjamin Spector. 2019. Distinctions between primary and secondary scalar implicatures. *Journal of Memory and Language* 106. 150–171. https://doi.org/10.1016/j.jml.2019.02.008.

Fox, Danny. 2007. Free choice and the theory of scalar implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and implicature in compositional semantics* (Palgrave Studies in Pragmatics, Language and Cog-

nition), 71–120. London: Palgrave Macmillan. https://doi.org/10.1057/97 80230210752_4.

Fox, Danny. 2014. Cancelling the maxim of quantity: Another challenge for a gricean theory of scalar implicatures. *Semantics and Pragmatics* 7(5). 1–20. https://doi.org/10.3765/sp.7.5.

Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29(5). 537–586. https://doi.org/10.1007/s1098 8-006-9004-4.

Fox, Danny & Benjamin Spector. 2018. Economy and embedded exhaustification. *Natural Language Semantics* 26(1). 1–50. https://doi.org/10.1007/s 11050-017-9139-6.

Frank, Michael C. & Noah D. Goodman. 2012. Predicting pragmatic reasoning in language games. *Science* 336(6084). 998. https://doi.org/10.1126/scie nce.1218633.

Franke, Michael. 2011. Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics and Pragmatics* 4(1). 1–82. https://doi.o rg/10.3765/sp.4.1.

Franke, Michael & Gerhard Jäger. 2016. Probabilistic pragmatics, or why Bayes' rule is probably important for pragmatics. *Zeitschrift für Sprachwissenschaft* 35(1). 3–44. https://doi.org/10.1515/zfs-2016-0002.

Frege, Gottlob. 1918. Der Gedanke: Eine logische Untersuchung. *Beitrage zur Philosophie des deutschen Idealismus* 1. 58–77.

Gajewski, Jon & Yael Sharvit. 2012. In defense of the grammatical approach to local implicatures. *Natural Language Semantics* 20(1). 31–57. https://d oi.org/10.1007/s11050-011-9074-x.

Gamut, L. T. F. 1991. *Language, logic and meaning*. Vol. 2. Chicago, IL: Chicago University Press.

Gazdar, Gerald. 1979. *Pragmatics: Implicature, presupposition, and logical form*. New York, NY: Academic Press.

Geurts, Bart. 2011. *Quantity implicatures*. Cambridge: Cambridge University Press. https://doi.org/doi:10.1017/CBO9780511975158.

Ginzburg, Jonathan. 1996. Dynamics and the semantics of dialogue. In Jerry Seligman & Dag Westerståhl (eds.), *Language, logic, and computation,* vol. 1 (CSLI Lecture Notes 58), 221–237. Stanford: Center for the Study of Language & Information (CSLI), Stanford University.

Goodman, Noah D. & Andreas Stuhlmüller. 2013. Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science* 5(1). 173–184. https://doi.org/10.1111/tops.12007.

Green, Georgia M. 1996. *Pragmatics and natural language understanding.* New York: Routledge.

Grice, Herbert Paul. 1975. Logic and conversation. In Peter Cole & Jerry Morgan (eds.), *Speech acts* (Syntax and Semantics 3), 41–58. Leiden: Brill. https://doi.org/10.1163/9789004368811_003.

Grice, Herbert Paul. 1989. *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Groenendijk, Jeroen & Floris Roelofsen. 2009. Inquisitive semantics and pragmatics. In Jesus M. Larrazabal & Larraitz Zubeldia (eds.), *Meaning, content, and argument: Proceedings of the ILCLI international workshop on semantics, pragmatics, and rhetoric*, 41–72. Leioa: University of the Basque Country Press.

Groenendijk, Jeroen & Martin Stokhof. 1984. *Studies on the semantics of questions and the pragmatics of answers.* Amsterdam: University of Amsterdam dissertation. https://hdl.handle.net/11245/1.392528.

Gutzmann, Daniel & Katharina Turgay. 2019. *Secondary content: The semantics and pragmatics of side issues.* Leiden: Brill. https://doi.org/10.1163/9789004393127.

Hara, Yurie. 2005. Contrastives and Gricean principles. In Paul Dekker & Michael Franke (eds.), *Amsterdam Colloquium 15*, 101–106. University of Amsterdam. https://archive.illc.uva.nl/AC/AC2005/uploaded_files/AC05Proceedings.pdf.

Harnish, Robert M. 1976. Logical form and implicature. In Thomas G. Bever, Jerrold J. Katz & Donald Terence Langendoen (eds.), *An integrated theory of linguistic ability.* New York, NY: Crowel.

Horn, Laurence R. 1972. *On the semantic properties of logical operators in english.* Los Angeles, CA: University of California, Los Angeles dissertation. https://www.proquest.com/docview/302676533.

Horn, Laurence R. 1984. Towards a new taxonomy of pragmatic inference: Q-based and R-based implicatures. In Deborah Schiffrin (ed.), *Meaning, form, and use in context*, 11–42. Georgetown University Press. https://repository.library.georgetown.edu/handle/10822/555477.

Horn, Laurence R. 1989. *A natural history of negation.* Chicago, IL: University of Chicago Press.

Horn, Laurence R. 2001. *A natural history of negation* (David Hume Series on Philosophy and Cognitive Science Reissues). Stanford, CA: CSLI Publications. First published in Horn 1989.

Hurford, James R. 1974. Exclusive or inclusive disjunction. *Foundations of Language* 11(3). 409–411. http://www.jstor.org/stable/25000785.

Katzir, Roni. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30(6). 669–690. https://doi.org/10.1007/s10988-008-9029-y.

Katzir, Roni & Raj Singh. 2014. Hurford disjunctions: Embedded exhaustification and structural economy. In Urtzi Etxeberria, Anamaria Fălăuş, Aritz Irurtzun & Bryan Leferman (eds.), *Sinn und Bedeutung 18*, 201–216. https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/313.

Leech, Geoffrey. 1983. *Principles of pragmatics*. London: Longman.

Matsumoto, Yo. 1995. The conversational conditions on Horn scales. *Linguistics and Philosophy* 18(1). 21–60. https://doi.org/10.1007/BF00984960.

Mayr, Clemens & Jacopo Romoli. 2016. A puzzle for theories of redundancy: Exhaustification, incrementality, and the notion of local context. *Semantics and Pragmatics* 9(7). https://doi.org/10.3765/sp.9.7.

Meyer, Marie-Christine. 2013. *Ignorance and grammar*. Cambridge, MA: Massachusetts Institute of Technology dissertation. http://hdl.handle.net/1721.1/84420.

Mill, John Stuart. 1865. *An examination of Sir William Hamilton's philosophy*. J.M. Robson (ed.) (The Collected Works of John Stuart Mill IX). Toronto: University of Toronto Press. http://www.jstor.org/stable/10.3138/j.ctt130jw1n.

Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In Kaarlo Jaakko Juhani Hintikka, Julius M.E. Moravcsik & Patrick Suppes (eds.), *Approaches to natural language* (Synthese Library 49), 221–242. Dordrecht: Springer. https://doi.org/10.1007/978-94-010-2506-5_10.

Pierrehumbert, Janet & Julia B. Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intensions in communication*, 271–311. Cambridge, MA: MIT Press. https://doi.org/10.7916/D8KD24FP.

Poortman, Eva B. 2017. *Concepts and plural predication: The effects of conceptual knowledge on the interpretation of reciprocal and conjunctive plural constructions*. Utrecht: Utrecht University dissertation. https://www.lotpublications.nl/concepts-and-plural-predication.

Potts, Christopher. 2013. *Conversational implicature: Interacting with grammar*. https://web.stanford.edu/~cgpotts/manuscripts/potts-interacting2013.pdf. Manuscript for the University of Michigan 2013 Workshop in Philosophy and Linguistics.

Recanati, François. 2001. What is said. *Synthese* 128(1-2). 75–91. https://doi.o rg/10.1023/A:1010383405105.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics, 1998 revision. In Jae Hak Yoon & Andreas Kathol (eds.), *Papers in semantics* (Working Papers in Linguistics 49). The Ohio State University, Department of Linguistics. A revised version was published in Roberts 2012.

Roberts, Craige. 2012. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5(6). 1–69. https://doi.org/10.3765/sp.5.6.

Roelofsen, Floris. 2013. A bare bone attentive semantics for *might*. In Maria Aloni, Michael Franke & Floris Roelofsen (eds.), *The dynamic, inquisitive, and visionary life of ϕ, ?ϕ, and $\diamond\phi$: A festschrift for Jeroen Groenendijk, Martin Stokhof, and Frank Veltman*, 190–215. ILLC publications. https://f estschriften.illc.uva.nl/Festschrift-JMF/papers/25_Roelofsen.pdf.

Roelofsen, Floris & Sam van Gool. 2010. Disjunctive questions, intonation, and highlighting. In Maria Aloni, Harald Bastiaanse, Tikitu de Jager & Katrin Schulz (eds.), *Logic, language and meaning: 17th Amsterdam colloquium, Amsterdam, the Netherlands, December 16-18, 2009, revised selected papers* (Lecture Notes in Computer Science 6042), 384–394. Berlin & Heidelberg: Springer. https://doi.org/10.1007/978-3-642-14287-1_39.

Rooth, Mats. 1992. A theory of Focus interpretation. *Natural Language Semantics* 1(1). 75–116. https://doi.org/10.1007/BF02342617.

Russell, Benjamin. 2006. Against grammatical computation of scalar implicatures. *Journal of Semantics* 23(4). 361–382. https://doi.org/10.1093/jos /ffl008.

Russell, Benjamin. 2012. *Probabilistic reasoning and the computation of scalar implicatures*. Providence, RI: Brown University dissertation. https://doi.o rg/10.7301/Z08S4N70.

Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27(3). 367–391. https://doi.org/10.1023/B:LING.00000233 78.71748.db.

Schulz, Katrin & Robert van Rooij. 2006. Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy* 29(2). 205–250. https://doi.org/10.1007/s10988-005-3760-4.

Scontras, Gregory & Michael Henry Tessler. 2018. Probabilistic language understanding: An introduction to the rational speech act framework. http s://www.problang.org (20 July, 2021).

Simons, Mandy. 2001. Disjunction and alternativeness. *Linguistics and Philosophy* 24(5). 597–619. https://doi.org/10.1023/A:1017597811833.

Simons, Mandy. 2011. Dynamic pragmatics, or why we shouldn't be afraid of embedded implicatures. *Semantics and Linguistic Theory (SALT)* 21. 609–633. https://doi.org/10.3765/salt.v21i0.2593.

Singh, Raj. 2008. On the interpretation of disjunction: Asymmetric, incremental, and eager for inconsistency. *Linguistics and Philosophy* 31(2). 245–260. https://doi.org/10.1007/s10988-008-9038-x.

Soames, Scott. 1982. How presuppositions are inherited: A solution to the projection problem. *Linguistic Inquiry* 13(3). 483–545. http://www.jstor.org/stable/4178288.

Spector, Benjamin. 2007. Scalar implicatures: Exhaustivity and Gricean reasoning. In Maria Aloni, Alastair Butler & Paul Dekker (eds.), *Questions in dynamic semantics* (Current Research in the Semantics / Pragmatics Interface 17), 225–250. Amsterdam & London: Elsevier. https://doi.org/10.1163/9780080470993_011.

van Tiel, Bob. 2014. *Quantity matters: Implicatures, typicality and truth*. Nijmegen: Radboud Universiteit Nijmegen dissertation. https://hdl.handle.net/2066/127889.

Uegaki, Wataru. 2015. *Interpreting questions under attitudes*. Cambridge, MA: Massachusetts Institute of Technology dissertation. http://hdl.handle.net/1721.1/99318.

Wagner, Michael. 2012. Contrastive topics decomposed. *Semantics and Pragmatics* 5(8). 1–54. https://doi.org/10.3765/sp.5.8.

Ward, Gregory & Julia Hirschberg. 1985. Implicating uncertainty: The pragmatics of fall-rise intonation. *Language* 61(4). 747–776. https://doi.org/10.2307/414489.

Westera, Matthijs. 2014. Exhaustivity through the maxim of relation. In Yukiko I. Nakano, Ken Satoh & Daisuke Bekki (eds.), *New frontiers in artificial intelligence. JSAI-isAI 2013* (Lecture Notes in Computer Science 8417), 193–205. Springer. https://doi.org/10.1007/978-3-319-10061-6_10.

Westera, Matthijs. 2017a. An attention-based explanation for some exhaustivity operators. *Sinn und Bedeutung* 21(2). https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/200.

Westera, Matthijs. 2017b. *Exhaustivity and intonation: A unified theory*. University of Amsterdam dissertation. https://eprints.illc.uva.nl/id/eprint/2143/.

Westera, Matthijs. 2017c. QUDs, brevity, and the asymmetry of alternatives. *Amsterdam Colloquium 21*. 502–510. http://hdl.handle.net/1887/316071 9.

Westera, Matthijs. 2018. Rising declaratives of the Quality-suspending kind. *Glossa: A journal of general linguistics* 3(1). 121. 1–32. https://doi.org/10 .5334/gjgl.415.

Westera, Matthijs. 2019. Rise-fall-rise as a marker of secondary QUDs. In Daniel Gutzmann & Katharina Turgay (eds.), *Secondary content: The semantics and pragmatics of side issues* (Current Research in the Semantics / Pragmatics Interface 37), 376–404. Leiden: Brill. https://doi.org/10.116 3/9789004393127_015.

Westera, Matthijs. 2020a. Hurford disjunctions: An in-depth comparison of the grammatical and the pragmatic approach. Manuscript. https://sema nticsarchive.net/Archive/jk5YzYyO/.

Westera, Matthijs. 2020b. Implying or implicating 'not both' in declaratives and interrogatives. *Sinn und Bedeutung* 24(2). https://doi.org/10.18148 /sub/2020.v24i2.906.

Zimmermann, Thomas Ede. 1989. Intensional logic and two-sorted type theory. *Journal of Symbolic Logic* 54(1). 65–77. https://doi.org/10.2307/2275 016.

Zimmermann, Thomas Ede. 2000. Free choice disjunction and epistemic possibility. *Natural Language Semantics* 8(4). 255–290. https://doi.org/10.1 023/A:1011255819284.

Matthijs Westera
Reuvensplaats 3-4
2311 BE Leiden
The Netherlands
m.westera@hum.leidenuniv.nl