

Relativized Exhaustivity: Mention-Some and Uniqueness

Yimei Xiang, Rutgers University

(Version: September 29, 2021)

Abstract *Wh*-questions with the modal verb *can* admit both mention-some (MS) and mention-all (MA) answers. This paper argues that we should treat MS as a grammatical phenomenon, primarily determined by the grammar of the *wh*-interrogative. I assume that MS and MA answers can be modeled using the same definition of answerhood (Fox 2013) and attribute the MS/MA ambiguity to structural variations within the question nucleus. The variations are: (i) the scope ambiguity of the higher-order *wh*-trace, and (ii) the absence/presence of an anti-exhaustification operator. However, treating MS answers as complete answers in this way contradicts the widely adopted analysis of uniqueness effects in questions of Dayal 1996, according to which the uniqueness effects of singular *which*-phrases arise from an exhaustivity presupposition, namely that a question must have a unique exhaustive true answer. To solve this dilemma, I propose that question interpretations presuppose ‘Relativized Exhaustivity’: roughly, the exhaustivity in questions is evaluated relative to the accessible worlds as opposed to the anchor/utterance world. Relativized Exhaustivity preserves the merits of Dayal’s exhaustivity presupposition while permitting MS; moreover, it explains the local-uniqueness effects in modalized singular *wh*-questions.

Keywords: interrogatives, questions, answers, mention-some, uniqueness, exhaustivity, exclusivity, free choice, modality, modal obviation, higher-order interpretations

Contents

1	Introduction	1
2	Distributional factors of MS	3
2.1	Modal flavor and modal force	3
2.2	Conversational goals	5
2.3	Agent dependency and conditional MS	6
3	Analytic directions	7
3.1	Pragmatic approaches	7
3.2	Semantic approaches: Nucleus-independent versus nucleus-dependent	9
3.2.1	Nucleus-independent approaches	10
3.2.2	Nucleus-dependent approaches	11
3.3	Arguments for nucleus-dependent approaches	13
3.3.1	The role of <i>can</i> in licensing MS in embeddings	14
3.3.2	The ‘mention-one-only’ constraint on MS answers	16
4	A nucleus-dependent approach to composing MS questions	19
4.1	General assumptions on questions and answers	19
4.1.1	Questions as topical properties	19

4.1.2	Answerhood	20
4.1.3	First-order versus higher-order interpretations	21
4.2	Deriving MS interpretations	22
4.2.1	Local exhaustification and first-order MS	22
4.2.2	More on the higher-order MS interpretation	23
4.2.3	Predictions on the distribution of MS	24
4.3	Deriving conjunctive MA interpretations	26
4.4	Deriving disjunctive MA interpretations	27
4.4.1	Two uses of <i>dou</i> : Exhaustivity marker and FC-trigger	28
4.4.2	Deriving the FC-trigger use: <i>dou</i> /DOU as an anti-exhaustification operator	30
4.4.3	Modal obviation of FC: Relativized Exclusivity	32
4.4.4	Deriving disjunctive MA	34
4.5	Interim summary	37
5	The dilemma	37
5.1	Dayal’s exhaustivity presupposition: merits and challenges	38
5.1.1	Uniqueness and singular <i>wh</i> -questions	38
5.1.2	Consequences of abandoning Dayal’s exhaustivity presupposition	39
5.1.3	A challenge to Dayal’s exhaustivity presupposition: local uniqueness	41
5.2	Two alternative accounts	42
5.2.1	Partition by exhaustification (Fox 2018, 2020)	42
5.2.2	Presuppositional <i>which</i> (Hirsch and Schwarz 2020)	45
6	Solving the dilemma: Relativized Exhaustivity	49
6.1	Relativized Exhaustivity	49
6.2	Predictions of RelExh on the distribution of MS	50
6.2.1	Permitting MS	50
6.2.2	Avoiding over-predictions of MS	52
6.3	Predictions of RelExh regarding uniqueness	52
6.3.1	Uniqueness effects in non-modalized questions	53
6.3.2	Uniqueness effects in <i>have to</i> -questions	53
6.3.3	Uniqueness effects in <i>can</i> -questions with a MS interpretation	54
6.3.4	Uniqueness effects in <i>can</i> -questions with a disjunctive MA interpretation	55
6.4	Interim summary	57
7	Conclusions	58
	Appendix A Choice questions versus MS questions	59
	Appendix B A variable-free treatment of modal bases	60

Relativized Exhaustivity: Mention-Some and Uniqueness

1. Introduction

Most questions call for a true answer that is exhaustive relative to the discourse domain. For example, to address the question in (1), the addressee A needs to specify all of the party attendants who are relevant to the interests of the questioner Q. Such answers are called *complete answers*.

- (1) (A's belief: Among the relevant individuals, only John and Mary went to the party.)
Q: 'Who went to the party?' A: 'John and Mary.'

If the addressee believes that she isn't fully informed and wants to be cooperative, she will mark the incompleteness of her answer explicitly. She may either say "I don't know who else did" or "I don't know if anyone else did", or utter the answer with a prosodic rise-fall-rise contour (indicated henceforth by '.../'; see Wagner et al. 2013).¹ Answers like (2a) are called *partial answers* or *incomplete answers*. If a partial answer is not properly marked, as in (2b), which has the default falling tone (indicated by '\'), it will give rise to an exclusivity inference and will be misleading to the questioner.

- (2) (A's belief: Antonio went to the party. It's unclear who else went to the party.)
Q: 'Who went to the party?'
A: a. 'Antonio did .../'
 L H* L-H%
 b. # 'Antonio did. \'
 L H* L-L% ↪ Only Antonio went to the party.

However, in many cases, *wh*-questions with the modal verb *can* (abbreviated as '*can*-questions') may be naturally addressed by a non-exhaustive answer. For instance in (3), the addressee A may felicitously choose to specify one of the accessible coffee places, as in (3a). Crucially, although this answer doesn't carry an ignorance mark, it doesn't give rise to an exclusivity inference. Following Groenendijk and Stokhof (1984), I call answers like (3a) '*mention-some* (MS) answers'. Relatedly, interpretations in which a question seeks a MS answer will be called '*MS interpretations*', and questions that admit MS interpretations will be called '*MS questions*'. To be sure, *can*-questions

¹There is no clear consensus on what the lack of final fall contributes to meaning. I take it to mark the pragmatic imperfection of an answer, roughly read as 'the best I can tell is ...'. For example, the answer in (i) is uttered without a final fall (marked by '...'). The lack of final fall indicates that A isn't sure whether her answer is relevant to the question, not that the answer is possibly non-exhaustive.

- (i) (A's belief: John went to the party. It's unclear whether he is a math professor.)
Q: 'Which math professor went to the party?' A: 'John did ...'

Moreover, answers that are semantically incomplete may have the default prosody if they are pragmatically optimal. In (ii), the non-exhaustive answer can be prosodically unmarked since it is informative enough relative to the questioner's conversational goals; indeed, answering the question exhaustively could sound like rambling.

- (ii) (To survive self-isolation, Alice stored a lot of food. She got lots of rice, canned chicken, and canned spinach — 100 lbs of each. She also got a variety of other things, such as dry noodles, canned tomatoes, and protein bars. Both Alice and her friend Bob hold the belief that one can survive by consuming a sufficient amount of carbs, protein, and fiber.)
Bob: 'What foods did you store to survive self-isolation?' Alice: 'Rice, canned chicken, and canned spinach. \'

also admit ‘*mention-all* (MA) interpretations’: in (3), A may address the question by listing all of the accessible coffee places. Hence, we say that *can*-questions exhibit a ‘MS/MA ambiguity’. MA answers to *can*-questions can be stated either as conjunctions as in (3b), or more naturally as disjunctions as in (3c) (Dayal 2017). The question interpretations in which a question is congruent with these two types of answers are called ‘conjunctive MA’ and ‘disjunctive MA’, respectively.

(3) (There are three coffee places nearby, namely Starbucks, Peet’s, and J.P. Licks.)

Q: ‘Where can we go to get coffee?’ / ‘Where can we get coffee?’

- A: a. ‘Starbucks.’ ↗ Starbucks is the only place to get coffee.
 b. ‘Starbucks, Peet’s, and J.P. Licks.’
 c. ‘Starbucks, Peet’s, or J.P. Licks.’

There are two directions that one can take in analyzing MS answers, namely, treating MS answers as partial answers or as complete answers. If one goes the first route, treating MS answers as partial answers, it is puzzling that *can*-questions systematically tolerate incompleteness. A common view on this puzzle is that whether a question admits a partial answer is primarily determined by pragmatic factors — in a goal-driven context, an answer that is semantically partial can be considered as complete relative to the conversational goals of the questioner. This view is taken by the ‘pragmatic approaches’, which consider MS as a simple pragmatic phenomenon (Sect. 3.1), as well as by most of the ‘semantic approaches’, which consider MS as an independent interpretation but attribute the licensing of MS to pragmatic factors (Sect. 3.2.1). These approaches typically pursue a joint analysis for MS answers to *can*-questions and non-exhaustive answers to non-*can*-questions.

There is no doubt that pragmatics plays an important role in the distribution of MS. What I want to address in this paper is the question whether pragmatics is, or can ever be, the primary source of MS. I observe that MS answers to *can*-questions are subject to a ‘mention-one-only’ constraint which cannot be explained by pragmatics: in response to a *can*-question, only the answers that specify exactly one option can be read non-exhaustively. Hence, contrary to most antecedent works but in line with George 2011: Chap. 6 and Fox 2013, this paper analyzes MS as a grammatical phenomenon that is primarily licensed by the presence of the modal verb *can*.

I will assume that MS answers and MA answers are derived based on a uniform answerhood operation which doesn’t require global exhaustivity (after Fox 2013). This operation encodes the properties of questions that relate to truth and exhaustivity/maximality. Next, I will present a compositional analysis that derives the MS/MA ambiguity based on structural variations within the question nucleus. In this analysis, the adopted answerhood operator delivers MS when it applies to certain forms of *can*-questions, and moreover, the yielded MS answers are ‘mention-one’, locally exhaustive, and mutually independent. I attribute the MS/MA ambiguity in *can*-questions to two structural variations: (i) the scope ambiguity of a higher-order *wh*-trace relative to *can*, and (ii) the absence/presence of an anti-exhaustification operator above *can*. As a welcome side effect of this analysis, parameter (ii) casts a bridge between the distribution of disjunctive MA and the modal obviation effect in licensing universal free-choice items.

However, allowing complete answers (‘complete’ in the sense that they are the expected type of direct answers to the respective questions) to be non-exhaustive conflicts with an influential exhaustivity presupposition from Dayal 1996: a question is defined only if it has a unique exhaustive true answer. This presupposition, henceforth called ‘*Dayal’s exhaustivity presupposition*’, nicely explains the uniqueness effects of singular *which*-phrases. To solve this dilemma, I will propose to replace Dayal’s exhaustivity presupposition with a presupposition of ‘*Relativized Exhaustivity*’, which has the

effect of evaluating exhaustivity relative to the accessible worlds as opposed to the anchor/utterance world. Relativized Exhaustivity permits MS where needed, without over-generating it. Moreover, it explains the local-uniqueness effects in modalized singular *wh*-questions.

The rest of this paper is organized as follows. Section 2 discusses the distributional factors of MS interpretations of matrix questions. Section 3 reviews existing approaches to the MS phenomenon and presents arguments for what I will call ‘nucleus-dependent approaches’, namely approaches that attribute the licensing of MS to factors that are dependent on the question nucleus. Section 4 compositionally derives the various interpretations of *can*-questions, including first-order/higher-order MS, conjunctive MA, and disjunctive MA. Section 5 delves into the dilemma between uniqueness and MS and reviews two recent analyses by Fox (2018, 2020) and Hirsch and Schwarz (2020). Section 6 proposes the concept of Relativized Exhaustivity, shows how it solves the dilemma, and accounts for local uniqueness. Section 7 concludes. The appendices discuss choice questions and explore a variable-free treatment of modal bases.

2. Distributional factors of MS

2.1. Modal flavor and modal force

Modal verbs express a quantification over a set of possible worlds accessible to the anchor world. Under normal matrix conditions, the anchor world is simply the utterance world. In the Kratzerian theory of modality, modal verbs vary along two axes, namely, *modal flavor* and *modal force*. In a modalized *wh*-question, the availability of MS is sensitive to both axes.

Modal flavor concerns how the possible worlds relate to the anchor world. It is jointly determined by *modal base* and *ordering source* (Kratzer 1981, 1991). Modals involved in a MS question are typically teleological or bouletic. These modals have a circumstantial modal base and an easy-to-perceive ordering source which provides a priority ranking related to someone’s goals or desires (Portner 2009). For example, the answer satisfying the MS question in (4) is read as: ‘Among the worlds compatible with the current circumstances, there is a world where our goals and desires are satisfied by us getting coffee at Starbucks.’

(4) Q: ‘Where can we get coffee?’ A: ‘Starbucks.’

In contrast, without contextual support, questions with an epistemic modal do not admit a MS interpretation (Dayal 2017: Chap. 3). In (5), the modal verb *could* quantifies over a set of worlds that are compatible with the available evidence. To properly answer the question, the addressee needs to list all the places that John possibly went to, as in (5b).

(5) (A’s belief: There are two coffee places near John’s home, namely, Starbucks and Peet’s. John frequents both.)

Q: ‘John left home for coffee 15 mins ago. Where could he have gone?’

A: a. ? ‘Starbucks.’

b. ‘Starbucks or Peet’s.’

Why can’t epistemic modals license MS? Intuitively, in (5) what the questioner is really interested in knowing is where John *actually* went, not where he *possibly* went. However, since the questioner doesn’t expect the addressee to know the answer, she chooses to ask a *could*-question to gather some possibilities. To maximize the chance that the answer covers the place that John actually went to, the addressee would list out all of the places he possibly went to.

Dayal (2017) has excluded epistemic modals and narrowed down the options to what Portner (2009) classifies as ‘priority modals’. In Portner’s classification, priority modals include not only the aforementioned teleological and bouletic modals but also deontic modals, whose ordering source is related to obligations and permissions. However, MS interpretations are difficult in questions with a deontic modal. In example (6), the modal verb *can* in the question is ambiguous between a teleological and a deontic flavor. If Bob intends to provide an easy path for Alice to get the job done, the MS answer (6a) is sufficient; however, if Bob intends to inform Alice about the regulations, perhaps for future reference, then the exhaustive answer (6b) is needed. The deontic reading is more salient if *can* is stressed or is replaced with *be allowed to*.²

- (6) (Alice, a first-year graduate student, is looking for someone to sign a document for her. According to the regulations, this document can be signed by either her mentor, the program director, or the department chair. One signature is sufficient. Bob knows the regulations well.)

Alice: ‘Who can I ask to sign this document?’

- Bob: a. ‘Your mentor.’ (Teleological: MS)
 b. ‘Your mentor, the program director, or the department chair.’ (Deontic: MA)

Felicitous answers to a deontic *can*-question must be exhaustive because for such questions, the addressee is expected to maximize the information of relevance to some pending decision of the questioner’s. For example, answer (6b) could be understood as follows: ‘You can ask your mentor to sign this document; however, if you can’t find your mentor or don’t want to interact with your mentor, you can ask the program director or the department chair to sign this document.’ This meaning would be an instance of what I call a ‘family of conditional MSs’ interpretation of *can*-questions, which is found in cases where the feasibility of an option varies situationally (see Sect. 2.3).

Modal force concerns the force of quantification, which can be existential, universal, or something in between. Only existential modals may license MS. In (7) and (8), the modal verb *should* is goal-oriented but has a universal modal force. In example (7), clearly, the addressee is expected to specify all the individuals who should be invited. Example (8) illustrates a multiple-choice scenario. If the choices are comparable, the addressee is expected to provide a free choice answer as in (8b). The single-choice answer (8a) is infelicitous or false in this context unless the addressee has reasons to prefer Starbucks over Peet’s (e.g., the Starbucks location is new). If the quantification domain of *should* is restricted to the worlds that best satisfy such preferences, (8a) becomes the only true answer.

- (7) Q: ‘Who should we invite to form a discussion panel?’
 A: ‘The department chair, the program director, and the two graduate representatives.’
- (8) (A’s belief: There are two coffee places nearby, namely, Starbucks and Peet’s.)
 Q: ‘Where should I get coffee?’
 A: a. ? ‘Starbucks.’
 b. ‘Starbucks or Peet’s. (Either is good.)’ ($\Box(\phi_{s.b.} \vee \phi_{p.ts.})$)

In sum, only teleological and bouletic modals with an existential force can license MS interpretations.³ In English, these modalities are realized as *can* or through the use of infinitives (as in *where to get coffee*). In this paper, unless specified, ‘*can*-questions’ refers to *wh*-questions with an existential teleological/bouletic modal.

²I thank Maria Biezma and Alexander Williams (pers. comm.) for helpful discussions on the data.

³Some works argue that existential indefinites also license MS. For arguments against this view, see Appendix A.

2.2. Conversational goals

Conversational goals play an important role in the distribution of MS and MA interpretations of matrix questions. On the one hand, questions without teleological/bouletic *can* admit MS answers in — and only in — goal-oriented contexts (Dayal 2017). (9) and (10) exemplify the case with the epistemic modal *might*. In both examples, the answer *Bill might be in* can be paraphrased as: ‘There is a world *w* compatible with the available evidence such that Bill is in the office in *w*.’ However, in (10), with an explicit conversational goal, this answer also implies a resolvedness inference: ‘In my belief, in the current circumstance, that Bill is in the office leads to a possible satisfaction of your goal to find someone to help you.’⁴ The same idea applies to the non-modalized question in (11).

- (9) a. I see a light on in the office. Who might be in at this time?
 b. Bill might be in.\
 ~→ *The available evidence only suggests that BILL might be in the office.*
- (10) a. I need help. Who might be in the office at this time?
 b. Bill might be in.\
 ↗ *The available evidence only suggests that BILL might be in the office.*
- (11) a. I need a ride to the party tonight. Who’s driving?
 b. Bill is.\
 ↗ *Only Bill is driving.* ((9)–(11) are modified from Dayal 2017: p. 77)

On the other hand, as exemplified in (12), a conversational goal that calls for an exhaustive answer sufficiently blocks the MS interpretation.

- (12) (After completing interviews with everyone on the long-list for a new faculty position, the departmental hiring committee decided to prioritize candidates who could teach Experimental Semantics or Field Methods.)
 Q: ‘Who can teach Experimental Semantics?’
 A: ‘Judy can.’
 ~→ *Among the candidates on the long-list, only Judy can teach Experimental Semantics.*

There are two directions to analyze the MS-licensing effect of goal-oriented contexts. One direction is to think of MS-licensing as a pragmatics-driven phenomenon: MS is primarily licensed by a non-exhaustive conversational goal, independent of the presence of a modal expression. In this view, the reason why the modal verb *can* licenses MS is that it may convey the goal-oriented modality. However, as I will lay out in Sect. 3.3, this ‘modal-to-context reduction’ analysis cannot explain the distribution of MS in embeddings and the ‘mention-one-only’ constraint.

The other direction, which I will pursue in this paper, is to take the modal verb *can* as the primary licensing factor; goal-oriented contexts may provide a constraint on the question nucleus (viz., the predicative part of a *wh*-interrogative that the *wh*-phrase combines with, which is typically analyzed as the IP part of the LF), forming a *can*-question which admits MS. For example, the questions in (10) and (11) can be paraphrased as follows, where the underlined parts are contributed by the context.

- (13) a. Which *x* is such that *x* might be in the office and that *x* can help me if s/he is in the office?

⁴Note that in (10), the literal meaning of the modalized answer, namely that Bill *might* be in the office, doesn’t ensure a possible satisfaction of the questioner’s goal. More precisely, this answer implicates that Bill could help you if he *is* in the office, not that Bill could help you if he *might* be in the office. Hence, no direct answer to *Who might be in the office?* (e.g., ‘Bill might be in the office’) ensures a possible satisfaction of the questioner’s goal.

- b. Which x is such that x is driving and that x *can* give me a ride if s/he is driving?

With this ‘context-to-modal reduction’ strategy, whatever explains the MS-licensing effect of the modal verb *can* also explains the MS-licensing effect of goal-oriented contexts. Moreover, this strategy predicts that, if a *wh*-construction cannot obtain a covert *can*-restriction from the context, it admits MS only in the presence of the modal verb *can*. As I will show, this prediction is consistent with the observations regarding question embeddings and non-interrogative *wh*-constructions, to be presented in Sect. 3.3.1.

2.3. Agent dependency and conditional MS

In some cases, MS answers are infelicitous even when the preceding question does have a goal-oriented *can* and an existential conversational goal. If the options are subject to agent-dependent restrictions (viz., the feasibility of an option varies with the desires of the questioner/agent or the situation that the questioner/agent is in) and the addressee is uncertain about these restrictions, she may want to specify multiple options that the questioner can choose from. In (14), although the questioner only wants to ship one package, the addressee provides two options to fulfill two different needs. In (15), although the questioner only needs one flu shot, the addressee lists all the options since she doesn’t know which option is convenient to the questioner. These cases are related to the observation made in Sect. 2.1 that deontic *can*-questions admit only exhaustive answers.

- (14) (There are four stores in town where one can ship packages: USPS, UPS, Fedex, and DHL. USPS has a cheaper rate, while the others provide express delivery. All these stores are open and are easily accessible to the questioner.)

Q: ‘Where can I go to ship this package?’

A: ‘USPS if you want a cheaper rate, or UPS if you need express delivery.’

- (15) (The city has locations of CVS, Walgreens, and Rite Aid, clustering in different districts. All of these stores provide flu vaccines.)

Q: ‘Where can I go to get a flu vaccine?’

A: ‘CVS, Walgreens, or Rite Aid, whichever of those is the most convenient to you.’

In the above two cases, the multiple-option answer addresses a family of conditional MS questions. For instance in (15), the question can be thought of as a family of sub-questions as follows: ‘If I live in x , where can I get a flu vaccine?’. In each such sub-question, the modal base of *can* is restricted to a set of worlds where the questioner lives in x .⁵ This meaning is formalized as below, where M is a circumstantial modal base, and \mathcal{C} is a set of propositional descriptions of relevant conditions:

$$(16) \quad \llbracket \text{Where can}_M \text{ I go to ...?} \rrbracket = \{ \llbracket \text{If } \phi, \text{ where can}_M \text{ I go to ...?} \rrbracket \mid \phi \in \mathcal{C} \} \\ = \{ \llbracket \text{Where can}_{M'} \text{ I go to ...?} \rrbracket \mid \exists \phi [\phi \in \mathcal{C} \wedge \forall w [M'_w = M_w \cap \phi]] \}$$

In each sub-question, the modal verb *can* is interpreted relative to a modal base M' which maps any anchor world w to a subset of M_w consisting of only the worlds where a condition in \mathcal{C} is true. A complete answer to this family of conditional *can*-questions should address all these questions, and therefore is the conjunction of a set of MS answers.

This paper will not delve into the ‘family of conditional MSs’ interpretations any further. However, for later discussions on the nature of the MS/MA distinction, it is crucial to distinguish between

⁵Following Kratzer (1981), I analyze the *if*-antecedent as a restriction of the modal base.

a single *can*-question with a MA interpretation and a family of conditional *can*-questions with a MS interpretation. To avoid confounding factors from the latter, the following discussion of the grammatical constraints on the MS/MA distinction will make the questioners’ conversational goals explicit and will consider only cases where the options are agent-independent.

3. Analytic directions

I classify existing approaches to MS as shown in Table 1. At the top level, *semantic approaches* differ from *pragmatic approaches* in that they consider MS an independent interpretation on a par with exhaustive interpretations, not purely a pragmatic phenomenon. In the view of semantic approaches, the MS/MA ambiguity is a genuine semantic ambiguity arising from the semantic composition of an interrogative sentence. Semantic approaches are further divided into two types, either *nucleus-independent* or *nucleus-dependent*, depending on whether or not they see the availability of MS as primarily determined by the question nucleus.

	Pragmatic	Semantic	
		Nucleus-independent	Nucleus-dependent
(i) Independence of meaning	No	Yes	Yes
(ii) Sensitivity to nucleus	No	No	Yes

Table 1: Pragmatic vs. semantic approaches. (i) *Independence of meaning*: MS exists as an independent interpretation of questions, on a par with the exhaustive interpretations; (ii) *sensitivity to nucleus*: the availability of MS is dependent on the question nucleus.

Pragmatic approaches highlight the importance of conversational goals in evaluating the adequacy of an answer; however, they pay less attention to the role of the modal verb *can* in the distribution of MS. Semantic approaches were first motivated to account for MS interpretations in question embeddings. Most semantic approaches are ‘nucleus-independent’. However, with regard to their predictions on what questions admit MS, nucleus-independent approaches are not any different from pragmatic approaches: nucleus-independent operations are independent of the form of the question nucleus; therefore, nucleus-independent approaches predict that any interrogative sentence is semantically ambiguous between MS and MA.

In contrast to pragmatic approaches and nucleus-independent approaches, nucleus-dependent approaches (e.g., George 2011: Chap. 6; Fox 2013) draw close attention to the connection between the availability of MS interpretations and the presence of the modal verb *can*.

3.1. Pragmatic approaches

MS was initially perceived as a pragmatic phenomenon. Earlier works adopted this view to maintain a core assumption in question semantics, namely, that question interpretations must be exhaustive (Groenendijk and Stokhof 1984).⁶ MS answers, which are non-exhaustive, were therefore treated as

⁶To be exact, Groenendijk and Stokhof (1984) assume that questions must be interpreted as *strongly exhaustive*, namely, the extensional meaning of an interrogative not only affirms all the true answers but also rules out all the false ones. In this view, the interrogative *who came*, for example, denotes a function that maps a world *w* to the exhausted proposition ‘only *x* came’ that is true in *w*. Hence for Groenendijk and Stokhof, the puzzle of MS is why certain questions admit interpretations that are not strongly exhaustive. Adopting this view, George (2011: Chap. 2) analyzes the MS/MA distinction as a semantic contrast between MS and strongly exhaustive interpretations (to be reviewed in Sect. 3.2.1).

In contrast to Groenendijk and Stokhof 1984, most recent works on question semantics take the so-called *weakly exhaustive* meaning (i.e., the meaning that only affirms all the true answers) as the basic meaning of an interrogative. For example, the

partial answers permitted due to pragmatic considerations, such as their being sufficient relative to the conversational goals. For instance, with the question *Where can we get coffee around here?*, if the goal of the questioner is just to get a coffee, the addressee is only expected to specify one coffee place suited to the questioner's needs; in contrast, if the goal is to investigate the local coffee market, the addressee is expected to list all the coffee places in the area under consideration. This direction of analysis was first sketched in partition semantics (Groenendijk and Stokhof 1984) and has remained popular in various frameworks of question semantics.

In contrast to Groenendijk and Stokhof 1984, a few other works on questions and answers assume that the semantics of an interrogative is underspecified, and further, that whether or not an answer is 'complete' is determined by how well this answer resolves the question relative to the goals of the questioner (Ginzburg 1995; van Rooij 2003, 2004; Schulz and van Rooij 2006; a.o.). Ginzburg (1995) argues that the *resolvedness* of an answer is evaluated relative to two context-dependent parameters: (i) the conversational goal(s), and (ii) the addressee's mental state, which determines whether an answer has the satisfaction of the goals as a consequence. For questions with a goal-oriented modal, resolvedness is straightforwardly satisfied. In (17), the goal 'give us a ride' is explicitly given by the question, and the true answer 'Bill can (give us a ride)' entails a possible satisfaction of the goal regardless of the addressee's mental state. This idea also applies to non-modalized questions. In (18), the answer 'Bill is (driving)' implies a resolvedness inference as follows: 'In my belief, in the current circumstance, that Bill is driving provides a way to satisfy your goal to find someone to give you a ride.'

(17) Q: 'Who can give us a ride?'

A: 'Bill can.'

(18) Q: 'I need a ride to the party tonight. Who's driving?'

A: 'Bill is. (He could give you a ride).'

van Rooij (2003, 2004) relates question-answering to *decision problems* and provides a statistical characterization for the *utility* of a non-exhaustive answer. In this theory, the fact that *can*-questions admit MS answers is not essentially different from the fact that the two non-*can*-questions in (19) admit non-exhaustive answers.

(19) a. Who has got a light? (with an existential goal)

b. Who, for example, came to the party? (with a partiality marker for *example*)

The most commonly raised challenge to pragmatic approaches, as pointed out by Groenendijk and Stokhof (1984) themselves and reiterated by George (2011), has been that MS interpretations are available not only in matrix questions but also in question embeddings. As seen in (20b), knowing a MS question entails knowing a true MS answer to this question.

(20) a. Jack knows [who arrived].

↪ For every relevant individual x , if x arrived, Jack knows that x arrived.

basic meaning of *who came* is a function that maps a world w to the proposition ' x came' such that it is true in w that only x came. Strong exhaustivity, then, is derived from weak exhaustivity via a separate answerhood operation (Heim 1994; Dayal 1996; Beck and Rullmann 1999) or a strengthening operation (Klinedinst and Rothschild 2011; Nicolae 2013, 2015; Uegaki 2015). Hence for these analyses, the puzzle of MS is why certain questions admit interpretations that are not even weakly exhaustive. Following this view, I will analyze the MS/MA distinction as a contrast between MS and weakly exhaustive interpretations.

- b. Jack knows [who can address this question].
 \rightsquigarrow For at least one relevant individual x such that x can address this question, Jack knows that x can address this question.

Pragmatic approaches have proposed multiple ways to address this challenge, by making either the semantic denotations of questions or the definition of answerhood context-dependent. For instance, as seen above, Ginzburg (1995) and van Rooij (2003) argue that the resolvedness of an answer is context-dependent, which can affect the truth conditions of question embeddings like (20b). Lahiri (2002) proposes that the interpretation of a question embedding involves picking a sub-question of the embedded question, whose size is determined by the goal of the speaker. An unpublished version of Dayal 2017 analyzes the semantic denotation of a question as context-independent but defines an answerhood operator for MS that is sensitive to the conversational goal. With these options, the fact that MS is available in embeddings isn't a knockdown argument against pragmatic approaches. However, it remains puzzling to pragmatic approaches as well as to some of the semantic approaches why MS is less readily available in embeddings of non-*can*-questions, to be discussed in Sect. 3.3.1.

3.2. Semantic approaches: Nucleus-independent versus nucleus-dependent

I call an approach to MS 'semantic' if it satisfies two criteria: (i) it perceives MS as an independent interpretation of questions, on a par with exhaustive interpretations, and (ii) it attributes the MS/MA ambiguity to operations in the semantic composition of the *wh*-construction. This classification isn't rigid. For example, as mentioned above, an unpublished version of Dayal 2017 assumes an answerhood operator for MS answers that involves context-dependent parameters. This approach can be viewed as 'semantic' because answerhood is part of the semantic composition, and it can be viewed as 'pragmatic' since the answerhood operator for MS answers is defined as context-dependent. However, as I will argue next, the demarcation between 'pragmatic' and 'semantic' isn't that important; what truly matters is whether the MS/MA ambiguity is attributed to operations that can interact with the modal verb *can*, which appears within the question nucleus.

Most existing semantic approaches are 'nucleus-independent approaches': they attribute the MS/MA ambiguity of a *wh*-question to operations *outside* the question nucleus (e.g., the answerhood operation). In contrast, nucleus-dependent approaches attribute this ambiguity to structural ambiguities *within* the question nucleus. Below I will review some representatives of both categories.

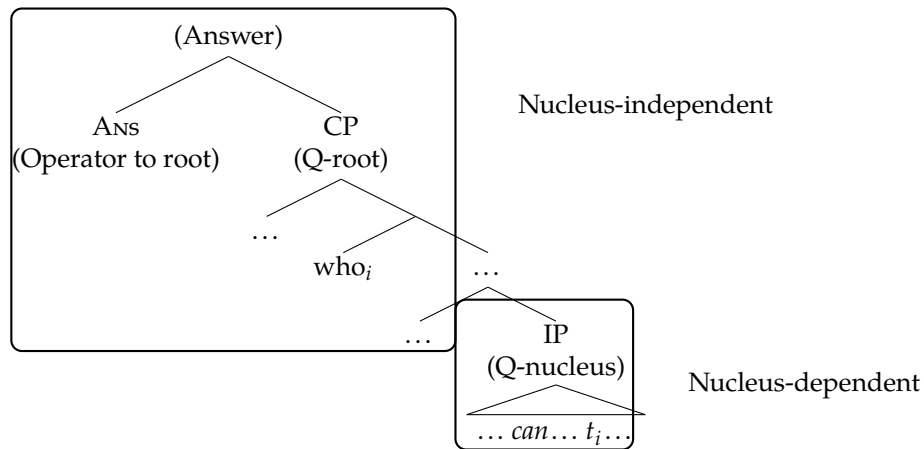


Figure 1: Nucleus-independent versus nucleus-dependent approaches

3.2.1. Nucleus-independent approaches

Nucleus-independent approaches attribute the MS/MA ambiguity to operations outside the question nucleus, such as the selection of a particular answerhood operator (Beck and Rullmann 1999; Caponigro and Davidson 2011), the lexical ambiguity of the *wh*-expression or the interrogative C head (Theiler et al. 2018), or the application of a strengthening/weakening operator outside the nucleus (George 2011: Chap. 2). The following reviews the accounts of Beck and Rullmann (1999) and George (2011: Chap. 2).

MS as existential answerhood Beck and Rullmann (1999) attribute the MS/MA ambiguity to the selection of one of several answerhood operators. They define the root of a question unambiguously as a Hamblin-Karttunen intension (i.e., a function that maps a world to the set of true propositional answers to the question in this world) but assume the availability of multiple answerhood operators, including:

- (21) a. $\llbracket \text{ANS}_{BR1} \rrbracket = \lambda w \lambda Q_{\langle s, \langle st, t \rangle \rangle} \cdot \bigcap \{ p \mid Q(w)(p) \wedge p(w) \}$
 b. $\llbracket \text{ANS}_{BR3} \rrbracket = \lambda w \lambda Q_{\langle s, \langle st, t \rangle \rangle} \lambda P_{\langle s, stt \rangle} \cdot \exists p [P(w)(p) \wedge Q(w)(p) \wedge p(w)]$

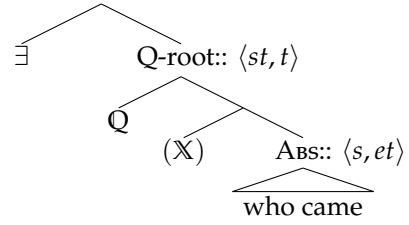
The above two answerhood operators differ in exhaustivity. In (22), for the embedding sentence ‘*x* knows *Q*’, applying ANS_{BR1} to the embedded question returns the conjunction of all true propositional answers to this question, yielding a MA (viz., weakly exhaustive) interpretation. In contrast, in (23), applying ANS_{BR3} to the embedded question returns an existential generalized quantifier over questions. Interpreting this quantifier over the embedding predicate yields an existential/MS interpretation.

- (22) a. LF for MA: $[x \text{ knows } [\text{ANS}_{BR1}(w)(Q)]]$
 b. Meaning: $\text{know}_w(x, \bigcap \{ p \mid Q(w)(p) \wedge p(w) \})$
 (23) a. LF for MS: $[\text{ANS}_{BR3}(w)(Q) \lambda p_{\langle s, t \rangle} \lambda w' [x \text{ knows}_{w'} p]]$
 b. Meaning: $\exists p [\text{know}_w(x, p) \wedge Q(w)(p) \wedge p(w)]$

MS as the absence of strong exhaustivity George (2011: Chap. 2) defines the concept of answerhood as unambiguously existential and attributes the MS/MA ambiguity to the absence/presence of a strengthening operator in question formation. As illustrated in (24), a question root is formed in two steps: (i) a property-forming abstraction operation Abs , and (ii) the shifting of this property into a set of propositions by a question-formation operator Q . In addition, before Q is applied, a strengthening operator X optionally acts on Abs . When X is absent, the root denotes a set of non-exhaustified propositions as in (24d), each of which is a non-strongly-exhaustive (viz., MS or weakly exhaustive) answer. When X is present, the root denotes a set of exhaustified propositions read as ‘Only the members of β came’ as in (24e), each of which is a strongly exhaustive answer.

(24) Who came?

- a. $\llbracket \text{Abs} \rrbracket = \lambda w \lambda x_e . \text{came}_w(x)$
- b. $\llbracket \text{Q} \rrbracket = \lambda \alpha_{\langle s, \tau t \rangle} \lambda p_{\langle s, t \rangle} . \exists \beta_\tau [p = \lambda w . \alpha(w)(\beta)]$
- c. $\llbracket \text{X} \rrbracket = \lambda \gamma_\tau \lambda \delta_\tau [\delta = \gamma]$
- d. Without X: MS/ weakly exhaustive
 $\llbracket \text{Q}(\text{Abs}) \rrbracket = \lambda p_{\langle s, t \rangle} . \exists \beta_e [p = \lambda w . \text{came}_w(\beta)]$
 $= \{ \lambda w . \text{came}_w(\beta) \mid \beta \in D_e \}$
- e. With X: strongly exhaustive
 $\llbracket \text{Q}(\text{X}(\text{Abs})) \rrbracket = \lambda p_{\langle s, t \rangle} . \exists \beta_{\langle e, t \rangle} [p = \lambda w [(\lambda x_e . \text{came}_w(x)) = \beta]]$
 $= \{ \lambda w . [(\lambda x_e . \text{came}_w(x)) = \beta] \mid \beta \in D_{\langle e, t \rangle} \}$



Since nucleus-independent approaches attribute the MS/MA ambiguity to operations independent of the question nucleus, they do not predict any structure-related constraints for the distribution of MS. For instance, no grammatical factor may block the use of Beck and Rullmann's ANS_{BR3} -operator or force the presence of George's X-operator. Hence, just like pragmatic approaches, nucleus-independent approaches predict that the distribution of MS is solely determined by pragmatics.

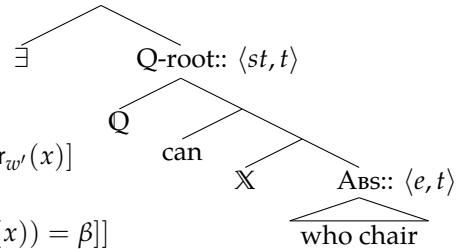
3.2.2. Nucleus-dependent approaches

Nucleus-dependent approaches attribute the MS/MA ambiguity to structural variations within the question nucleus. Below I review the accounts of George (2011: Chap. 6) and Fox (2013), which analyze the MS/MA ambiguity in *can*-questions as a scope ambiguity of exhaustivity or distributivity relative to the modal *can*.

MS as a scopal effect of exhaustification Distinct from the analysis reviewed in (24), George (2011: Chap. 6) assumes that the X-operator is mandatorily used in question formation and treats MS as a scopal effect of this operator. When the X-operator takes scope below an existential expression, such as the existential modal *can*, the root denotes a set of propositions that are not globally exhaustive.

(25) Who can chair the committee?

- a. $\llbracket \text{Abs} \rrbracket = \lambda x_e . \text{chair}_{w'}(x)$
- b. $\llbracket \text{X}(\text{Abs}) \rrbracket = \lambda \delta_{\langle e, t \rangle} [\delta = \lambda x_e . \text{chair}_{w'}(x)]$
- c. $\llbracket \text{can} \rrbracket = \lambda q_{\langle s, t \rangle} . \exists w' \in M_w [q(w')]$
- d. $\llbracket \text{can}(\text{X}(\text{Abs})) \rrbracket = \lambda \delta_{\langle e, t \rangle} . \exists w' \in M_w [\delta = \lambda x_e . \text{chair}_{w'}(x)]$
- e. $\llbracket \text{Q}(\text{can}(\text{X}(\text{Abs}))) \rrbracket$
 $= \lambda p_{\langle s, t \rangle} . \exists \beta_{\langle e, t \rangle} [p = \lambda w . \exists w' \in M_w [(\lambda x_e . \text{chair}_{w'}(x)) = \beta]]$
 $= \{ \lambda w . \exists w' \in M_w [(\lambda x_e . \text{chair}_{w'}(x)) = \beta] \mid \beta \in D_{\langle e, t \rangle} \}$



This account predicts that the MS interpretation is only available in questions with an existential expression. It also nicely accounts for what I call the 'local-exhaustivity effect' of MS answers, an observation made in Xiang 2016: Chap. 2. Compare the answers in (26): although it is true that Andy can serve on the committee, (26a) is a bad answer, in contrast to (26b,c), each of which specifies the full composition of a possible committee. This contrast argues that MS answers are subject to local exhaustivity; for example, (26b) is read as: 'It can be the case that *only* Andy and Billy serve on the committee.' In George's scope-based account, the strengthening operator X applied under *can*

captures this intuition.

- (26) (The committee can be formed in two ways: it should either have the two members Andy and Billy, or have the three members Andy, Billy, and Cindy.)

Who can serve on the committee?

- a. # Andy can.\
- b. Andy and Billy can.\
- c. Andy, Billy, and Cindy can.\

However, this account also faces several problems. First, the composition between *can* and $\mathbb{X}(\text{Abs})$ doesn't follow the conventional rules: *can* selects for a proposition, while $\mathbb{X}(\text{Abs})$ denotes a set of predicates. Second, although the local application of \mathbb{X} nicely predicts the local-exhaustivity effect of MS answers, the global application of \mathbb{X} forces a strongly exhaustive interpretation, which is too strong (see fn. 6). Third, this account predicts that any expressions with an existential force, including indefinites, can license MS. However, a number of empirical distinctions between *can*-questions and questions with an existential indefinite argue that the non-exhaustive interpretations of these two types of questions have different origins. For details, see Appendix A and references therein.

MS as a scopal effect of distributivity Fox (2013) assumes that MS and MA answers are obtained by a uniform answerhood operator which calls for complete true answers but doesn't demand global exhaustivity. As defined in (27), a true answer is complete as long as it is not asymmetrically entailed by any other true answers. Here Q denotes the Hamblin set of the question, also called 'answer space'.

$$(27) \text{ANS}_{\text{Fox}}(w)(Q) = \{p \mid w \in p \in Q \wedge \forall q[w \in q \in Q \rightarrow q \not\subseteq p]\} \quad (\text{Fox 2013})$$

Fox's answerhood operator is stronger than simple existentiality but weaker than exhaustivity. In contrast, Dayal's (1996) answerhood operator returns the exhaustive true answer, which is the unique true answer that entails all the true answers.⁷

$$(28) \text{ANS}_{\text{Dayal}}(w)(Q) = \exists p[w \in p \in Q \wedge \forall q[w \in q \in Q \rightarrow p \subseteq q]].$$

$$i p[w \in p \in Q \wedge \forall q[w \in q \in Q \rightarrow p \subseteq q]] \quad (\text{Dayal 1996})$$

Fox (2013) further analyzes the MS/MA ambiguity as a scopal effect of distributivity. He assumes that the *wh*-trace, together with a covert distributivity operator, forms a distributive phrase $[X \text{ EACH}]$. If this phrase takes scope below the modal verb *can*, as in (29a), the answer space of the question is not closed under conjunction. In this interpretation, applying ANS_{Fox} may return a set of non-exhaustive/MS answers. In contrast, if distributivity takes wide scope or if the question nucleus contains no existential expression, the answer space of the question is closed under conjunction, and the output of applying ANS_{Fox} is a singleton set containing only the conjunctive MA answer.

⁷The following illustrates the difference between Dayal's and Fox's answerhood operators:

- (i) Let $Q_1 = \{p, q\}$ and $Q_2 = \{p, q, p \wedge q\}$, where p and q are propositions that are true in w and logically independent of each other. We have:
 - a. $\text{ANS}_{\text{Dayal}}(w)(Q_1)$ is undefined, while $\text{ANS}_{\text{Fox}}(w)(Q_1) = \{p, q\}$;
 - b. $\text{ANS}_{\text{Dayal}}(w)(Q_2) = p \wedge q$, while $\text{ANS}_{\text{Fox}}(w)(Q_2) = \{p \wedge q\}$.

Fox's answerhood operator predicts the following distribution of MS: a question with an answer space Q has a MS interpretation only if there is a world w such that $\text{ANS}_{\text{Fox}}(w)(Q)$ is a non-singleton set (as in the case of Q_1). If Q is closed under conjunction (as in the case of Q_2) or if the propositions in Q are all mutually exclusive, $\text{ANS}_{\text{Fox}}(w)(Q)$ remains a singleton set containing only the proposition yielded by $\text{ANS}_{\text{Dayal}}$, and in this case the question has a MA interpretation.

(29) (Only Andy and Billy can chair the committee. Co-chairing is disallowed.)

Who can chair the committee?

- a. $[_{CP} \text{ who } \lambda X \dots [_{IP} \text{ can } [[X \text{ EACH}] \lambda x [x \text{ chair the committee }]]]]]$ (*can* \gg *EACH*: MS)
 i. True answers: $\{\diamond \text{chair}(a), \diamond \text{chair}(b)\}$
 ii. $\text{ANS}_{\text{Fox}}(w)(Q) = \{\diamond \text{chair}(a), \diamond \text{chair}(b)\}$
- b. $[_{CP} \text{ who } \lambda X \dots [_{IP} [X \text{ EACH}] \lambda x [\text{ can } [x \text{ chair the committee }]]]]]$ (*EACH* \gg *can*: MA)
 i. True answers: $\{\diamond \text{chair}(a) \wedge \diamond \text{chair}(b), \diamond \text{chair}(a), \diamond \text{chair}(b)\}$
 ii. $\text{ANS}_{\text{Fox}}(w)(Q) = \{\diamond \text{chair}(a) \wedge \diamond \text{chair}(b)\}$

Compared with George’s account, Fox’s account allows for weak exhaustivity and doesn’t force strong exhaustivity. However, it doesn’t explain the local-exhaustivity effect of MS answers and under-generates MS answers. For the case in (26), repeated below, the MS answer (26b) won’t count as a complete true answer since it is asymmetrically entailed by (26c). In other words, Fox’s account predicts that (26b) is as bad as (26a), contrary to fact.

(26) (The committee can be formed in two ways: it should either have the two members Andy and Billy, or have the three members Andy, Billy, and Cindy.)

Who can serve on the committee?

- a. # Andy can. \ $\diamond \text{serve}(a)$
 b. Andy and Billy can. \ $\diamond(\text{serve}(a) \wedge \text{serve}(b))$
 c. Andy, Billy, and Cindy can. \ $\diamond(\text{serve}(a) \wedge \text{serve}(b) \wedge \text{serve}(c))$

In addition, just like George’s account, Fox’s account over-predicts a MS-licensing effect for existential indefinites. This over-prediction problem also applies to the recently developed partition-by-exhaustification account of Fox 2018, 2020, to be reviewed in Sect. 5.2.1.

In short, nucleus-dependent approaches attribute the origin of the MS/MA ambiguity to operations within the question nucleus. Since the modal verb *can* is interpreted within the nucleus, only nucleus-dependent approaches may predict a grammatical relation between the presence of *can* and the availability of MS.

3.3. Arguments for nucleus-dependent approaches

Pragmatic approaches and nucleus-independent approaches predict that the MS interpretations of *can*-questions are primarily licensed by pragmatic factors, especially by whether or not the question has an existential conversational goal. In this view, the modal verb *can* may license MS just because it naturally comes with an existential conversational goal. This idea is compatible with the data on matrix questions, as seen in (9)–(11). Further, as for why existential conversational goals license MS, pragmatic approaches attribute this fact to the resolvedness/utility of an answer: when the conversational goal is existential, a non-exhaustive answer suffices for resolving the question. To this extent, as predicted by the utility theory of van Rooij (2003, 2004), the non-exhaustive interpretations of *can*-questions and non-*can*-questions (see (19)) are of the same nature.

In contrast, nucleus-dependent approaches predict that the modal verb *can* (or another equivalent modal expression), which appears within the question nucleus, is the primary source for licensing MS. Existing nucleus-dependent approaches haven’t said much about the MS-licensing effect of

conversational goals. However, they are compatible with the context-to-modal reduction strategy which I argued for in Sect. 2.2: conversational goals may license MS for matrix non-*can*-questions simply because they can restrict the question nucleus and turn a non-*can*-question into a question with a covert *can*. Other pragmatic factors may serve as a blocker for MS answers or a rescuer for incomplete answers.

The following discusses two linguistic constraints on MS, both of which support the direction of nucleus-dependent approaches. First, without the presence of *can*, MS interpretations are difficult to obtain in question embeddings and non-interrogative *wh*-constructions. Second, MS answers to *can*-questions are subject to a ‘mention-one-only’ constraint, which cannot be explained by pragmatic factors.

3.3.1. The role of *can* in licensing MS in embeddings

In contrast to matrix questions, MS interpretations are less readily available in embeddings of a non-*can*-question (Dayal 2017: Chap. 3). This observation has been experimentally validated by Xiang and Cremers (2017): all else being equal, the presence of *can* in the embedded question significantly increases the acceptance of a MS interpretation.

The experiments of Xiang and Cremers (2017) proceeded as follows. First, the participants were presented with a paragraph describing the background:

“Mary is in charge of choosing two children to lead the dance. The only rule is that the children leading the dance should have an accessory in common.”

This background specifies an existential goal of the attitude holder (i.e., Mary wants to find two children who have an accessory in common to co-lead the dance). Next, the participants saw a set of pictures illustrating how children were dressed and two sentences describing Mary’s memory. Figure 2 presents a MS trial: Mary remembers only one of the two pairs of children who have an accessory in common and can co-lead the dance.

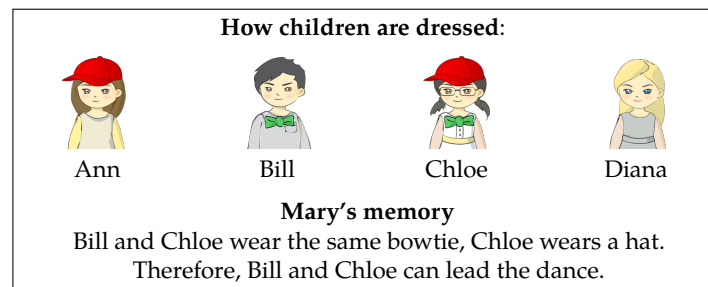


Figure 2: MS trial in Xiang and Cremers 2017

Finally, the participants were asked to judge the truth value of a question-embedding sentence. The form of the embedded question varied by two conditions: (i) whether the *wh*-subject is *who* or *which children*, and (ii) whether the embedded question is a *can*-question, as in (30a), or is not modalized, as in (30b). With the background described above, whatever addresses the embedded *can*-question in (30a) also addresses the embedded non-modalized question in (30b), and vice versa.

- (30) a. Mary remembers {who, which children} can lead the dance. ([+MODAL])
 b. Mary remembers {who, which children} have an accessory in common. ([-MODAL])

This study found that statements with *can* were judged acceptable significantly more often than ones without. In the statistical analysis, regardless of the form of the *wh*-subject, fitting the data of the MS trials with a logistic mixed-effects model yielded a significant effect of [\pm MODAL] ($p < .001$). (For detailed results, see Xiang and Cremers 2017.)

Other than with question embeddings, the MS-licensing effect of *can* is also observed in non-interrogative *wh*-constructions such as *wh*- free relatives and Mandarin *wh*-conditionals. These constructions have an existential interpretation only in the presence of an existential teleological/bouletic modal. *Wh*- free relatives usually have a universal interpretation; however, in (31b), the free relative that contains the existential goal-oriented modal *could* accepts and even prefers an existential interpretation (Chierchia and Caponigro 2013). In (31c), with the universal goal-oriented modal *ought to*, the free relative has to be interpreted universally.

- (31) a. Jack ate [what Mary cooked for him].
 \rightsquigarrow *John ate everything that Mary cooked for him.* (Universal)
- b. Jack went to [where he could go to get help].
 \rightsquigarrow *John went to one of the places where he could go to get help.* (Existential)
- c. Jack went to [where he ought to go to get help].
 \rightsquigarrow *John went to all of the places where he ought to go to get help.* (Universal)

Wh-conditionals in Mandarin are made up of two *wh*-clauses with the same *wh*-morphology. In most cases, a *wh*-conditional expresses a universal condition: every entity that constitutes a true short answer to the question expressed by the antecedent *wh*-clause also constitutes a true short answer to the question expressed by the consequent *wh*-clause.⁸ However, as seen in (32b), a *wh*-conditional has an existential interpretation if the antecedent *wh*-clause contains the existential goal-oriented modal verb *neng* ‘can’ (Liu 2016b; Xiang 2016, 2021a).

- (32) a. Ni qu-guo nar, wo jiu qu nar.
 you go-EXP where, I jiu go where
 Intended: ‘I will go to every place where you have been to.’ (Universal)
- b. Nar neng mai-dao jiu, wo jiu qu nar.
 where can buy-reach liquor, I jiu go where
 Intended: ‘I will go to one of the places where I can buy liquor.’ (Existential)

In sum, despite the fact that existential conversational goals sufficiently license non-exhaustive answers to matrix questions, in a number of embedding constructions, including question embeddings, *wh*- free relatives, and *wh*-conditionals, MS/existential interpretations are less readily available in the absence of an existential goal-oriented modal like *can*. One plausible explanation for this contrast is that the contextual support for licensing MS or other non-exhaustive interpretations is less accessible to embedded *wh*-constructions. If this explanation is on the right track, we can conclude that the goal-oriented modal *can* is a sufficient MS-licenser.⁹

⁸For definitions of ‘short answers’, see fn. 14.

⁹There is a possible counterargument against the view of treating *can* as the primary MS-licenser that I would like to address. Dayal (2017: Sect. 3.2) observes that the interrogative subject of *depend on* does not have a MS interpretation, as in (i). She then argues that MS requires a goal-driven modality, which is unavailable in sentences with a non-human subject. This idea predicts that the licenser of MS isn’t the expression *can*, but rather the non-exhaustive goal introduced by *can*.

(i) Where you can get gas depends on what day it is.

I would argue that the interrogative subject in (i) is semantically ambiguous between MS and MA, but this ambiguity collapses

3.3.2. The ‘mention-one-only’ constraint on MS answers

Distinct from other non-exhaustive answers, MS answers to *can*-questions are subject to a ‘mention-one-only’ constraint: a felicitous MS answer only specifies one option that resolves the question. Therefore, it is more precise to call MS answers ‘*mention-one* answers’ — in contrast to ‘*mention-few* answers’, which specify multiple options. In what follows, I will argue that *mention-few* differs from *mention-one* in two aspects: (i) embeddings of *can*-questions admit *mention-one* interpretations but not *mention-few* interpretations; (ii) in discourse, unlike *mention-one* answers, *mention-few* answers to *can*-questions easily imply exclusivity.

First, embeddings of *can*-questions allow for *mention-one* and *mention-all* interpretations but not non-exhaustive ‘*mention-N*’ ($N \geq 2$) interpretations, even in cases where *mention-N* fits better with the conversational goal than *mention-one/all*. In (33), the conversational goal calls for a ‘*mention-three* answer’ (i.e., an answer that specifies three possible venues). If MS interpretations were primarily licensed by the conversational goal, (33) should have the *mention-three* interpretation (33b), contrary to fact.

- (33) (The committee needs to identify three possible campus venues for the upcoming conference. They expect about 300 people to attend the conference.)

Jill knows where on campus one can hold a conference with 300 participants.

- a. ✓ ‘Jill knows one/all of the places on campus where one can hold a conference with 300 participants.’ (Available: *mention-one/all*)
 b. ✗ ‘Jill knows three of the places on campus where one can hold a conference with 300 participants.’ (Unavailable: *mention-three*)

Note that it is difficult to rule out *mention-few* interpretations based on the truth conditions of (33). The sentence ‘*x* knows *Q*’, where *Q* has a MS interpretation, is true if and only if (i) *x* knows a true *mention-one* answer to *Q*, and (ii) *x* has no false belief relevant to *Q*. However, since knowing a true *mention-three* answer or knowing three *mention-one* answers entails condition (i), the *mention-one* interpretation is true in a scenario where Jill knows three possible venues. To examine the availability of a *mention-three* interpretation empirically, consider the following conversation involving a polar question with embedded *can*. In (34), despite the conversational goal being ‘*mention-three*’, the addressee Bob cannot felicitously reply with a denial or an apology while admitting that he knows one possible venue. Such a reply would be felicitous only if Alice had explicitly requested him to name three places, e.g., if she had asked “Do you know/ Could you tell me three places on campus

under the selectional requirement of the matrix predicate. The subject of *depend on* must be strongly exhaustive. For example in (ii), if the subject interrogative were weakly/intermediately exhaustive, speaker B’s objection would be infelicitous.

- (ii) A: ‘Who got selected depended on who passed the exam.’
 B: ‘Well, that’s not true. Bill was selected even though he didn’t pass the exam.’

The strong exhaustivity inference can be derived from a MS denotation in the same way in which it is derived from a weakly exhaustive denotation. As shown in (iii), this inference can be defined in terms of answerhood equivalence (Heim 1994; Dayal 1996; Beck and Rullmann 1999): For a question *Q* and an evaluation world *w*, the strong exhaustivity inference is the set of worlds *w'* such that the weakly exhaustive answers to *Q* in *w* and *w'* are the same. Crucially, this definition applies even if *Q* has a MS interpretation: let $\text{ANS}(\llbracket Q \rrbracket)(w)$ be the set of true MS answers to *Q* in *w*; then (iii) denotes the set of worlds *w'* such that the MS answers to *Q* in *w* and in *w'* are the same, which corresponds to the strong exhaustivity inference.

- (iii) Strong exhaustivity inference of a question *Q* in the world *w*: $\lambda w' [\text{ANS}(w)(\llbracket Q \rrbracket) = \text{ANS}(w')(\llbracket Q \rrbracket)]$

Hence, it is plausible that in (i) the subject interrogative itself has a MS interpretation, but thanks to the selectional constraint of *depend on*, this MS semantic denotation is turned into a strong exhaustivity inference by a question-external operation.

where one can hold a conference with 300 participants?”, or if she had added to her original question: “Could you name three such places?”¹⁰

(34) Alice: ‘We are looking for a campus venue for the upcoming conference. We need to identify three options. Do you know where on campus one can hold a conference with 300 participants?’

Bob: ‘Yes, but I only know one such place.’/‘#No/Sorry, I only know one such place.’

Second, in discourse, in contrast to mention-one answers, mention-few answers are not read non-exhaustively in the absence of an ignorance marker. The following is a comprehension task concerning a typical MS question, presented in English and in Mandarin. Both tests were posted informally to social network sites, given in writing and thus free of prosodic biases. Speakers were asked to judge whether either the mention-one answer (i) or the mention-two answer (ii) gave rise to an exclusivity inference. Notably, among those who judged (i) non-exhaustively, a large majority judged (ii) exhaustively.¹¹

(35) (Alice has an electric slicer, which comes with 10 blades. These blades have different colors and shapes, designed for different ingredients. Now, while Alice is cooking, her friend Bob comes to help her cut carrots. Bob thinks highly of Alice’s cooking skills and trusts her words.)

a. *English*

Bob: Which blade can I use to cut carrots?

Alice: i. The green one.

ii. The green one or the black one.

b. *Mandarin*

Bob: Na-ge daopian keyi yonglai qie huluobo?
which-cl blade can be-used-for cut carrot?

‘Which blade can be used for cutting carrots?’

¹⁰Note that in this scenario it is infelicitous to ask “Do you know *which three places* on campus one can hold a conference with 300 participants?”, because this question presupposes that there are only three such places on campus. This inference derives from the uniqueness effects of numeral-modified *which*-phrases (see Sect. 5.1.1).

¹¹The design of the comprehension task in (35) avoids a few confounding factors. First, the domain of the *wh*-phrase is small and salient, which avoids confounds from implicit domain restrictions. Admittedly, a small domain makes the exhaustive interpretations more appealing, which makes it harder to gather judgments related to the MS interpretation. Second, this example avoids the ‘family of conditional MSs’ interpretation (discussed in Sect. 2.3), which is otherwise hard to distinguish from mention-few/all.

The participants were asked to reply ‘Yes’/‘No’ to the following: (Q1) If you were Bob, when you hear Alice says (i), would you be inclined to believe that only the green blade is suited for cutting carrots? (Q2) If you were Bob, when you hear Alice says (ii), would you be inclined to believe that only the green blade and the black blade are suited for cutting carrots?

For the English version, (35a), most responses collected from native speakers were unrelated to the MS interpretations. Out of 18 consultants, 13 replied ‘Yes’ to both Q1 and Q2 (abbreviated as 1Y2Y), which shows that they interpreted the question exhaustively. In particular, 5 of these 13 speakers reported that the exclusivity inference was stronger in (ii) than in (i), which supports the ‘mention-one-only’ constraint. Among those who read the question non-exhaustively, 3 replied 1N2Y, which again affirms the constraint, and 2 replied 1N2N based on considerations related to implicit restrictions on the modal base: (i) implies that the green blade is the only choice fitting Alice’s preferences, but in principle there could be other blades suited for cutting carrots. To better validate the claimed ‘mention-one-only’ constraint, it would be helpful to conduct a more extensive experimental study in English.

For the Mandarin version, (35b), the pattern is clearer, thanks to the larger size of the data set. Out of 606 native speakers, 308 interpreted the question exhaustively. Among the 251 speakers who read (i) non-exhaustively, 214 interpreted (ii) exhaustively, which well supports the ‘mention-one-only’ constraint. The distribution of the votes was as follows: 1Y2Y (308), 1N2N (37), 1Y2N (47), 1N2Y (214).

- Alice: i. Lüse-de zhe-ge keyi.
green-MOD this-CL can
‘The green one can.’
- ii. Lüse-de zhe-ge huozhe heise-de zhe-ge dou keyi.
green-MOD this-CL or black-MOD this-CL DOU can
Intended (free choice): ‘The green one and the black one both can.’

The contrast in exhaustivity between (i) and (ii) argues that there is a grammatical constraint that only allows mention-one answers to be MS answers. As I will argue in Sect. 4.2, when a *can*-question has a MS interpretation, answers expressed by an (atomic/plural) individual are possibly complete, while answers formed out of a Boolean coordination are not.

For answers to questions with a partiality marker (e.g., *for example, for instance, give me an example*), there is no contrast in exhaustivity between mention-one and mention-few. In (36) and (37), regardless of the presence of *can*, neither mention-one nor mention-few answers imply exclusivity. What’s more, a questioner using a partiality marker has the option of making an exact ‘mention-*N*’ inquiry, by explicitly saying “Give me *N* examples”.

- (36) Who is on your committee, for example?
a. Andy is on my committee. \ ↗ Only Andy is on my committee.
b. Andy and Billy are on my committee. \ ↗ Only Andy and Billy are on my committee.
- (37) Who can chair the committee alone, for example?
a. Andy can. \ ↗ Only Andy can chair it alone.
b. Andy and Billy (each) can. \ ↗ Only Andy and Billy can chair it alone.

The contrast between the *can*-question (35) and the *for example*-questions (36),(37) in accepting non-exhaustive mention-few answers argues that the sources of non-exhaustivity in these two types of questions are different. I treat the partiality marker *for example* as a discourse-level expression which appears outside the question root. It signals that the questioner tolerates an incomplete true answer and presupposes the existence of such an answer in the answer space of the question (Xiang 2021a).¹² In contrast, the non-exhaustivity of MS answers to *can*-questions is grammatically obtained from the

¹²I define the presupposition of *for example* as follows:

- (i) $[[Q, \textit{for example?}]]$ is defined in w only if there is a proposition p such that
a. p is a complete true answer to Q in some other world w' (viz., p is potentially complete);
b. p is asymmetrically entailed by a complete true answer to Q in w (viz., p is true but partial in w).

In brief, *for example* can be used only if the preceding question has answers that are both potentially complete and possibly partial. This treatment is supported by the distributional constraint observed in (ii), taken from Xiang 2021a: *for example* cannot be felicitously attached to a question that has at most one true answer.

- (ii) a. Which boy came, # for example?
b. Is it raining, # for example?
c. Did you vote for Andy or Billy, # for example?

Note that for question (37), my analysis of MS questions predicts that the presupposition of *for example* can only ever be satisfied if and only if the preceding *can*-question has a MA interpretation. In this interpretation, the MA answer is the unique complete true answer, and the mention-one/few answers in (37a,b) are incomplete answers asymmetrically entailed by the MA answer. In contrast, the presupposition of *for example* can never be satisfied if the preceding *can*-question has a MS interpretation. In a MS interpretation, the question has no potentially complete answer that can ever be incomplete: the complete true answers are all mention-one answers, and the mention-one answers, unless independently ruled out, are all potentially complete.

answerhood operator for complete true answers. In Fox’s account and my account, the answerhood operator may return non-exhaustivity when it applies to particular types of root denotations of *can*-questions.

In sum, MS answers to *can*-questions are subject to a ‘mention-one-only’ constraint. For one thing, the sentence ‘*x* knows *can*-Q’ implies that *x* knows a mention-one answer or the mention-all answer to the embedded *can*-question, but not that *x* knows a non-exhaustive mention-few answer to this *can*-question. For another, to serve as replies to a *can*-question, non-exhaustive mention-few answers, while being more informative than mention-one answers, have to be ignorance-marked like partial answers. This ‘mention-one-only’ constraint cannot be explained by pragmatics (for my own explanation, see Sect. 4.2).

Finally, let me clarify a possible point of confusion: ‘mention-one’ refers to mentioning one option/possibility, not one atomic individual/entity. In contrast to the above examples, where each option is constituted by one single individual/location/entity, there are also cases where an option is made up of the sum or a Boolean coordination of multiple individuals/locations/entities. For example, to answer (26) *Who can serve on the committee?*, a mention-one answer should specify the sum of a group of individuals who can simultaneously serve on the committee. In (38), the two answers (a) and (b), which are presented as conjunctive sentences, are mention-one answers. They each specify a way to group teams and assign leaders, and they each name a Boolean conjunction over individuals.¹³

- (38) (The players can be grouped into 2–3 teams. Each team needs one or two leaders.)
- Who can we ask to lead a team?
- a. We can ask Alex to lead a team and Ben to lead a team. (A two-teams option)
 - b. We can ask Alex to lead a team, Ben to lead a team, and Clark and David together to lead a team. (A three-teams option)

4. A nucleus-dependent approach to composing MS questions

This section will first lay out the relevant background assumptions on question semantics (Sect. 4.1). Next, in what constitutes a core part of this paper, I will propose a nucleus-dependent approach to composing *can*-questions. The proposal will cover a variety of interpretations of *can*-questions, including first-order MS and higher-order MS (Sect. 4.2), conjunctive MA (Sect. 4.3), and disjunctive MA (Sect. 4.4).

In particular, the treatment of MS offered here will account for three semantic properties of MS answers, namely local exhaustivity, mutual independence, and ‘mention-one-only’. The treatment of MA relates disjunctive MA to the derivation and the modal obviation effect of universal free choice.

4.1. General assumptions on questions and answers

4.1.1. Questions as topical properties

In line with categorial approaches, I assume that the semantic denotations of questions are *topical properties* (after Chierchia and Caponigro 2013). As exemplified in (39a,b), the topical property of a *wh*-question is a function that maps an individual in the *wh*-domain to a proposition in the answer

¹³The context of (38) might make the MA interpretation more salient. However, what matters here is that the answers (a) and (b) are both mention-one answers, not mention-two/three.

space. In other words, as formalized in (39c), the answer space of a question is the *image* (i.e., the set of all output values) of the topical property of this question. I henceforth denote a topical property as ‘[[Q]]’ and an answer space as ‘Q’.

- (39) Q: ‘Which books did John read?’ A: ‘The Harry Potter books.’
- a. $[[Q]] = \lambda x_e : x \in \text{books}_@.\lambda w[\text{read}_w(j, x)]$
 - b. $[[Q]]([[HP]]) = \text{hp} \in \text{books}_@.\lambda w[\text{read}_w(j, \text{hp})]$
 - c. $Q = \{[[Q]](x) \mid x \in \text{Dom}([[Q]])\}$
 $= \{\lambda w.\text{read}_w(j, x) \mid x \in \text{books}_@\}$

Defining questions as topical properties, rather than as partitions of possible worlds or sets of propositions, makes it easy to track short answers and makes the analysis of MS applicable to the existential readings of *wh*-free relatives and *wh*-conditionals.¹⁴ Moreover, this analytical choice is crucial for my solution to the dilemma between uniqueness and MS: I will argue that question interpretation is subject to a condition called ‘Relativized Exhaustivity’, the definition of which requires the extraction of short answers from question denotations. The main reason for this choice is that, in a modalized *wh*-question, the semantics of a propositional answer varies by modal base, while that of a short answer doesn’t (for a detailed explanation, see fn. 40 in Sect. 6). However, see also Appendix B for a variable-free analysis of modal bases which avoids this issue.¹⁵

4.1.2. Answerhood

In contrast to the classic theories of question semantics (Karttunen 1977; Groenendijk and Stokhof 1984; a.o.), many theories developed in the past three decades (Heim 1994; Dayal 1996; Beck and Rullmann 1999; Fox 2013; a.o.) encode the question properties of truth and maximality/exhaustivity in answerhood operators, not in the root denotations of questions. An answerhood operator applies to the evaluation world and the root denotation of a question and returns the (set of) true answer(s) that fulfill the assumed requirement of maximality/exhaustivity. Some of these more recent analyses (e.g., Beck and Rullmann 1999) assume multiple answerhood operators to account for the variations in exhaustivity in question interpretations.

I assume that there is only one type of maximality, henceforth called ‘*maximal (max-)informativity*’, that is grammatically encoded in answerhood operators. I adopt the concept of answerhood in Fox 2013: a true answer is complete if and only if it is not asymmetrically entailed by any other true answers. Such an answer will be called a ‘*max-informative true answer*’, in contrast to what Dayal’s (1996) answerhood operator yields as the unique ‘*strongest/exhaustive true answer*’. As mentioned in the review of Fox 2013 in Sect. 3.2.2, max-informativity is stronger than simple existentiality but weaker than exhaustivity. It derives either MS or MA, depending on the logical relation of the propositions in the answer space of the question.¹⁶

¹⁴The term ‘short answers’ usually refers to the linguistic representations in a question–answer discourse that specify only the new information of an answer. Short answers in discourse are either analyzed as bare nominals (i.e., they are simple DPs denoting meanings in the *wh*-domain), or are treated as covertly clausal (i.e., they are clauses with ellipsis and are interpreted as propositions). For this paper, whether a short answer should be analyzed as a bare DP or as the elliptical form of a clause isn’t important. I use the term ‘short answers’ mainly to refer to the meanings in the *wh*-domain. When *wh*-questions are interpreted as topical properties, such meanings are the possible arguments of the function-like denotation of a *wh*-question.

¹⁵This paper stays neutral on how exactly topical properties are compositionally derived. I will only specify the composition of the question nucleus. Besides categorial approaches (Hausser and Zaefferer 1979; Hausser 1983; Xiang 2021a) and the variants called ‘structured-meaning approaches’ (von Stechow and Zimmermann 1984; Ginzburg and Sag 2000; Krifka 2001a), dynamic approaches (Dotlačil and Roelofsen 2019, 2021; Haoze Li 2019, 2021) also allow for extraction of short answers from question denotations.

¹⁶Here the MA interpretation refers to the weakly exhaustive interpretation. As for the contrast between weak exhaustivity

Adapting the ANS_{Fox} -operator in (27) to the assumed question semantics, I define two answerhood operators as follows, where these operators are applied to a topical property, not a Hamblin set. The superscripts S and P stand for ‘short’ and ‘propositional’, respectively. $\llbracket \text{Q} \rrbracket_w$ is the abbreviation of the set of short answers to Q that are true in w (formally: $\llbracket \text{Q} \rrbracket_w := \{\alpha \mid \alpha \in \text{Dom}(\llbracket \text{Q} \rrbracket) \wedge w \in \llbracket \text{Q} \rrbracket(\alpha)\}$).

(40) Answerhood operators (to be modified)

a. For complete true short answers:

$$\text{ANS}^S(w)(\llbracket \text{Q} \rrbracket) = \{\alpha \mid \alpha \in \llbracket \text{Q} \rrbracket_w \wedge \forall \beta [\beta \in \llbracket \text{Q} \rrbracket_w \rightarrow \llbracket \text{Q} \rrbracket(\beta) \not\subset \llbracket \text{Q} \rrbracket(\alpha)]\}$$

b. For complete true propositional answers:

$$\text{ANS}^P(w)(\llbracket \text{Q} \rrbracket) = \{\llbracket \text{Q} \rrbracket(\alpha) \mid \alpha \in \text{ANS}^S(w)(\llbracket \text{Q} \rrbracket)\}$$

This definition does not consider conditions needed for explaining uniqueness effects. Modifications will be made in Sect. 6.

4.1.3. First-order versus higher-order interpretations

Wh-questions are semantically ambiguous between first-order and higher-order interpretations (Spector 2007, 2008; Xiang 2021b). For example, the following *wh*-question, containing a universal modal, can be completely addressed by specifying some particular books, as in (41a), or by using a generalized quantifier (GQ) over a set of books, as in (41b). (The example is taken from Spector 2007.)

(41) Which books does John have to read?

a. The French novels.

b. The French novels or the Russian novels. (The choice is up to him.) ($\square \gg or$)

I assume the following LFs and topical properties for the question in (41):

(42) First-order interpretation: ‘For which x , such that x is a plurality of books, is it the case that John has to read x ?’

a. $[_{CP} \text{ which-books } \lambda x_e [_{IP} \text{ have-to } [_{VP} \text{ John read } x]]]$

b. $\llbracket \text{Q} \rrbracket = \lambda x_e : x \in \text{books}_{@} . \square \lambda w [\text{read}_w(j, x)]$

(43) Higher-order interpretation: ‘For which π , such that π is a GQ over books, is it the case that John has to read π ?’

a. $[_{CP} \text{ which-books } \lambda \pi_{\langle et, t \rangle} [_{IP} \text{ have-to } [\pi \lambda x_e [_{VP} \text{ John read } x]]]]$

b. $\llbracket \text{Q} \rrbracket = \lambda \pi_{\langle et, t \rangle} : \pi \in {}^H\text{books}_{@} . \square \lambda w [\pi(\lambda x_e . \text{read}_w(j, x))]$

In LF (43a), the fronted *wh*-phrase binds a higher-order trace π across the modal verb (after Spector 2007, 2008). The interactions between the higher-order *wh*-trace and other scopal expressions inside the question nucleus offer scope ambiguities for free; thus, for a scope-based account of MS, there is no need to assume an additional scopal expression inside the question nucleus — in contrast to the analyses of George (2011) and Fox (2013), which analyze MS as a scopal effect of exhaustivity or distributivity. In denotation (43b), ‘ ${}^H\text{books}_{@}$ ’ stands for a set of GQs ranging over a set of entities that are books in the actual world @. Although not all GQs can serve as semantic answers to *wh*-questions

and strong exhaustivity, I assume that strong exhaustivity is derived by the application of a partitioning operation. As argued in fn. 9, for any evaluation world w , the strong exhaustivity inference of a question Q is simply $\lambda w' [\text{ANS}(w)(\llbracket \text{Q} \rrbracket) = \text{ANS}(w')(\llbracket \text{Q} \rrbracket)]$, where ‘ANS’ could refer to either the ANS^S -operator in (40a) or the ANS^P -operator in (40b). This partitioning operation is viewed as an answerhood operator in many existing works (Heim 1994; Dayal 1996; Beck and Rullmann 1999), but it could also be viewed as an operation external to question denotations and answerhood operations.

(Spector 2007, 2008; Xiang 2021b), for this paper, all that matters is that the domain of a higher-order topical property includes Montagovian individuals and their Boolean coordinations.

4.2. Deriving MS interpretations

4.2.1. Local exhaustification and first-order MS

I assume that the first-order MS interpretation of a *can*-question is derived based on the following LF. The core assumption is that an exhaustification operator O (\approx *only*) (Chierchia et al. 2012; a.o.) is applied to the local VP and is associated with the individual *wh*-trace x (of type e).

- (44) Who can serve on the committee? (First-order MS)
 $[_{CP} \text{ who } \lambda x_e [_{IP} \text{ can } [_{OC} [_{VP} x_{[+v]} \text{ serve on the committee }]]]]$

The local O -operator is assumed to account for the local exhaustivity and mutual independence of MS answers. In (45), repeated from (26), local exhaustivity says that a MS answer should specify all the members of a possible committee: although it is true that Andy can serve on the committee, (45a) is not a good MS answer. Mutual independence is a property predicted by Fox’s definition of answerhood: by Fox’s (2013) definition, (45b) counts a good MS answer only if it isn’t asymmetrically entailed by (45c).

- (45) (The committee can be formed in two ways: it should either have the two members Andy and Billy, or have the three members Andy, Billy, and Cindy.) (= (26))

Who can serve on the committee?

- a. # Andy can. \
- b. Andy and Billy can. \
- c. Andy, Billy, and Cindy can. \

As defined in (46), the O -operator affirms the prejacent proposition and negates the alternatives of the prejacent that are not entailed by that prejacent (Chierchia et al. 2012; a.o.). The domain variable C carried by the O -operator denotes a contextually determined subset of the alternatives.

- (46) $[[O_C]] = \lambda p \lambda w. p(w) = 1 \wedge \forall q \in C [p \not\subseteq q \rightarrow q(w) = 0]$

Inserting an O -operator under the modal verb *can* captures the aforementioned two properties of MS answers: the O -operator asserts local exhaustivity and makes the individual answers mutually independent. This consequence is similar to what is achieved by George’s locally applied \mathbb{X} -operator (see (24)); however, as demonstrated below, the O -operator is technically neater than the \mathbb{X} -operator in composition.

The LF (44) is now computed as follows. Here the O -operator is associated with an e -type *wh*-trace x . The $[+v]$ feature of the trace activates a set of ‘variable alternatives’, defined like focus alternatives as in (48).¹⁷ The domain variable C carried by the O -operator denotes a contextually determined subset of the variable alternatives of the VP.¹⁸ Composing this LF yields the topical property (47b).

¹⁷I assume a separate feature $[+v]$ (cf. the focus feature $[+F]$) because variable alternatives are specific to variable-denoting expressions and do not require focus marking.

¹⁸The association relation between an exhaustivity operator and a c -commanded expression is realized via the following *domain restriction condition* (Xiang 2020), which expands on the focus association condition of Rooth 1996: For any operator Θ quantifying over a domain C and combining with an expression δ , if Θ agrees with an alternative-activating feature $[+x]$, $[[\Theta_C(\delta)]]$ is defined only if $C \subseteq x\text{-ALT}(\delta)$. This condition ensures that in (47) the domain variable C carried by the O -operator denotes a subset of $v\text{-ALT}(VP)$.

In the context described in (45), applying the assumed answerhood operator to this topical property returns a set consisting of two max-informative true answers, given in (47c), each of which is a MS answer.

- (47) $[\text{CP who } \lambda x_e [\text{IP can } [O_C [\text{VP } x_{[+v]} \text{ serve on the committee }]]]]$ (= (44))
- $\text{v-ALT}(\text{VP}) = \{\phi_x \mid x \in D_e\}$ (ϕ_x abbreviates ‘ x serves on the committee’)
 - $[\text{Q}] = \lambda x_e : x \in \text{hmn}_@ \cdot \diamond O_C \phi_x$, where $C \subseteq \{\phi_x \mid x \in \text{hmn}_@\}$
 - $\text{ANS}^P(w)([\text{Q}]) = \{\diamond O_C \phi_{a \oplus b}, \diamond O_C \phi_{a \oplus b \oplus c}\}$
- (48) Variable alternatives: For any trace and pronoun α ,
- $$\text{v-ALT}(\alpha) = \begin{cases} D_{\text{type}([\alpha])} & \text{if } \alpha \text{ carries a } [+v] \text{ feature} \\ \{[\alpha]\} & \text{otherwise} \end{cases}$$

I assume that the insertion of the local O -operator is preferred but not mandatory. As observed in (49), the local exhaustivity implicature is cancellable and suspendable. This observation also argues that the presence of the local O -operator should only affect cancellable inferences such as local exhaustivity; in contrast, mandatory effects such as the uniqueness effects of singular *which*-phrases and the unavailability of MS in non-*can*-questions should be independent of the presence of this O -operator. Section 6 will account for these mandatory effects without resorting to a local O -operator.

- (49) Who can serve on the committee?
- Andy and Billy can ... maybe also Cindy.
 - Andy and Billy can. I don’t know whether we should add a third person.

4.2.2. More on the higher-order MS interpretation

The derivation of the higher-order MS interpretation is illustrated in (50). Here the example sentence uses the predicate *chair the committee* (same as in (29)), which avoids complications from plural answers. Compared with the derivation of the first-order MS interpretation in (47), the only difference is that here the *wh*-phrase undertakes an IP-internal movement before reaching [Spec, CP], which creates a higher-order trace π (of type $\langle et, t \rangle$) between the modal verb *can* and the local O -operator. Composing this LF yields the higher-order topical property (50b), where ‘ $\text{hmn}_@$ ’ stands for a set of GQs over human individuals.

- (50) Who can chair the committee? (Higher-order MS)
- $$[\text{CP who } \lambda \pi_{\langle et, t \rangle} [\text{IP can } [\pi \lambda x_e [O_C [\text{VP } x_{[+v]} \text{ chair the committee }]]]]]]$$
- $\text{v-ALT}(\text{VP}) = \{\phi_x \mid x \in D_e\}$ (ϕ_x abbreviates ‘ x chairs the committee’)
 - $[\text{Q}] = \lambda \pi_{\langle et, t \rangle} : \pi \in \text{hmn}_@ \cdot \diamond \pi(\lambda x_e \cdot O_C \phi_x)$, where $C \subseteq \{\phi_x \mid x \in \text{hmn}_@\}$

Figure 3 illustrates the answer space yielded by the LF in (50). This illustration contains four answers related to the two individuals Andy (a) and Billy (b), derived by applying the topical property (50b) to the Montagovian individuals a^\uparrow and b^\uparrow , their Boolean conjunction $a^\uparrow \cap b^\uparrow$, and their Boolean disjunction $a^\uparrow \cup b^\uparrow$.¹⁹ For example, $\diamond O_C \phi_a$ is derived based on a^\uparrow and is read as ‘There is a world w compatible with the current circumstances such that only a chairs the committee in w ’.

¹⁹For any meaning α of type τ , we have: the Montague-lifted meaning of α is α^\uparrow (of type $\langle \tau t, t \rangle$) such that $\alpha^\uparrow := \lambda m_{\langle \tau, t \rangle} \cdot m(\alpha)$. Boolean conjunctions and disjunctions are defined in terms of set intersection and union, respectively. For any meanings α and β of type τ , we have: $\alpha^\uparrow \cap \beta^\uparrow := \lambda m_{\langle \tau, t \rangle} \cdot m(\alpha) \wedge m(\beta)$, and $\alpha^\uparrow \cup \beta^\uparrow := \lambda m_{\langle \tau, t \rangle} \cdot m(\alpha) \vee m(\beta)$.

(Only Andy and Billy can chair the committee. Co-chairing is disallowed.)

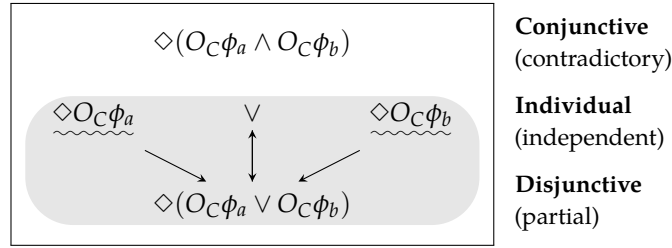


Figure 3: The answer space of the higher-order MS interpretation of *Who can chair the committee?* [Arrows indicate entailment relations, shading marks the true answers, and underwaving marks the max-informative true answers. The bi-implication between ‘ \vee ’ in the middle and the disjunctive answer at the bottom means that this disjunctive answer is logically equivalent to the disjunction of the two individual answers.]

The semantic properties of the three types of answers can be characterized as follows:

- *Conjunctive answers*: Due to the exclusivity of the local O -operator, all conjunctive answers are contradictory. For example, in Figure 3, $\diamond(O_C\phi_a \wedge O_C\phi_b)$ is read as ‘#There is a world w compatible with the current circumstances such that only a chairs the committee in w and only b chairs the committee in w ’.
- *Individual answers*: The individual answers can be true and are logically independent of each other. Moreover, due to the non-monotonicity of the O -operator, mutual independence also applies to plural answers (e.g., $\diamond O_C\phi_{a \oplus b}$), no matter whether the predicate *chair the committee* is read as distributive (i.e., ‘chair the committee separately’) or collective (i.e., ‘chair the committee together’). Hence, any true individual answer is a max-informative true answer.
- *Disjunctive answers*: In Figure 3, the disjunctive answer $\diamond(O_C\phi_a \vee O_C\phi_b)$ is logically equivalent to the disjunction of the two individual answers ($\diamond O_C\phi_a$ and $\diamond O_C\phi_b$). Due to this equivalence, whenever the disjunctive answer is true, one of the individual answers is true and asymmetrically entails the disjunctive answer. Hence, in the higher-order MS interpretation, disjunctive answers can never be max-informative and are always partial.

This analysis explains why MS answers to *can*-questions are all ‘mention-one’. In responding to a *can*-question with a MS interpretation, only individual answers, each of which specifies one single option, are possibly max-informative. In this case, the addressee cannot use a Boolean coordination to express a non-exhaustive mention-few answer: conjunctive answers convey a contradiction, and disjunctive answers are partial answers implicating epistemic ignorance. Hence, if the addressee uses a Boolean conjunction or disjunction to convey a mention-few answer, she must be understanding the question with a MA interpretation, which calls for an exhaustive answer. This mention-few answer, if it is non-exhaustive, should be ignorance-marked just like any partial answer.

To sum up, by adopting Fox’s definition of answerhood and assuming an exhaustification operator under the existential modal *can*, this analysis explains three semantic properties of MS answers to *can*-questions, namely local exhaustivity, mutual independence, and ‘mention-one-only’.

4.2.3. Predictions on the distribution of MS

According to Fox’s definition of answerhood, a question has a MS interpretation if and only if there is a world in which the answer space of this question has multiple max-informative true propositions.

This requirement is met only if the two conditions below are met:

- (A) the semantically independent answers are not mutually exclusive;
- (B) the answer space is not closed under conjunction.

In a *can*-question, the presence of the existential modal *can* allows for the satisfaction of both conditions. For condition (A), the local *O*-operator makes the individual answers logically independent, and further, the presence of an existential modal above the *O*-operator ensures that these answers are not mutually exclusive and allows multiple individual answers to be simultaneously true. In comparison, with other conditions being equal, if the existential modal is dropped as in (51b) or replaced with a universal modal as in (51c), the individual answers are mutually exclusive.

- (51) Let $C = \{\phi_x \mid x \in D\}$. For any a and b in D such that $\phi_a \neq \phi_b$, we have:
- a. $\diamond O_C \phi_a \wedge \diamond O_C \phi_b \not\Rightarrow \perp$
 - b. $O_C \phi_a \wedge O_C \phi_b \Leftrightarrow \perp$
 - c. $\square O_C \phi_a \wedge \square O_C \phi_b \Leftrightarrow \perp$

For condition (B), the answer space of a *can*-question with a higher-order interpretation is not closed under conjunction if the higher-order *wh*-trace takes scope below the existential modal. In Figure 3 above, the answer space does not contain a proposition equivalent to the conjunction of the two individual answers. In contrast, for a *wh*-question without a modal or with a universal modal like *have to*, the answer space derived in a higher-order interpretation is closed under conjunction: in each illustration in Figures 4a–c below, the conjunctive answer is semantically equivalent to the conjunction of the two individual answers (indicated by the bi-implication between ‘ \wedge ’ in the middle and the conjunctive answer at the top).

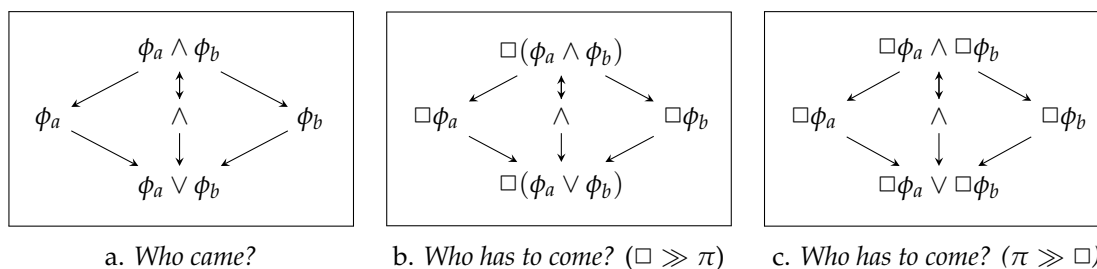


Figure 4: Answer spaces of non-modalized questions and *have to*-questions

However, the assumptions made in this section cannot fully explain why MS interpretations are *only* available in *can*-questions: it’s possible to satisfy conditions (A) and (B) even if the question doesn’t contain an existential modal. For a concrete example, consider the non-modalized question in (52). Here the predicate *form a team* is ‘stubbornly collective’ (Xiang 2021b): in contrast to *form teams*, *form a team* doesn’t admit a covered/(semi-)distributive interpretation. Although the answer space derived in the higher-order interpretation is closed under conjunction, the true answers allowed in the first-order interpretation are logically independent, as seen in (52a). Hence, unless we find independent reasons to rule out (52a), the assumptions made in this section would predict a MS/MA ambiguity for (52), contrary to fact. (Q_w abbreviates the set of true propositional answers in w .)

- (52) (The children formed two teams: $a + b$ formed one, and $c + d$ formed the other.)
Which children formed a team?

- a. First-order interpretation:
 ‘For which x , such that x is a plurality of children, is it the case that x formed a team?’
 $Q_w = \{\phi_{a\oplus b}, \phi_{c\oplus d}\}$ (ϕ_x abbreviates ‘ x formed a team’)
- b. Higher-order interpretation:
 ‘For which π , such that π is a GQ over children, is it the case that π formed a team?’
 $Q_w = \{\phi_{a\oplus b}, \phi_{c\oplus d}, \phi_{a\oplus b} \wedge \phi_{c\oplus d}, \phi_{a\oplus b} \vee \phi_{c\oplus d}\}$

There are two ways to solve the MS over-generation problem in (52). One way is to enrich the answer space of the first-order interpretation so that it is closed under conjunction. Fox (2018, 2020) assumes that *wh*-phrases may quantify over higher-order pluralities. According to this assumption, the conjunctive answer ‘ $a + b$ formed a team, and $c + d$ formed a team’ is derived based on the higher-order plurality $\{\{a, b\}, \{c, d\}\}$ (for details, see fn. 33 in Sect. 5.2.1). In this proposal, if higher-order pluralities are available in the first-order *wh*-quantification, the answer space of (52a) is closed under conjunction.

The other way to solve the problem is to rule out (52a) by an independent constraint. As I will argue in Sect. 5.1.2, the MS over-generation problem arises not only in *wh*-questions with a non-distributive predicate, but also in questions with a uniqueness effect (e.g., *Which boy came?*) and *wh*-questions with an existential indefinite (e.g., *Which movie(s) did one of the boys watch?*). It would be more appealing to rule out these over-generating cases uniformly.

Note that conceptually the over-generation problem laid out in this section is the price we pay for abandoning Dayal’s exhaustivity presupposition: if question interpretations had to be exhaustive, any interpretation that does not render the existence of an exhaustive true answer would be deviant. To predict the distribution of MS, I will argue for a condition called ‘Relativized Exhaustivity’, which works like Dayal’s exhaustivity presupposition except in modalized *wh*-questions. This condition makes several welcome predictions (Sect. 6.2), one of which is that only *wh*-questions with an existential modal can have an interpretation that allows for multiple max-informative true answers.

4.3. Deriving conjunctive MA interpretations

Recall that the MA answer to a *can*-question can be expressed either as a conjunction or as a disjunction. I argue that the two forms of MA answers correspond to two distinct MA interpretations of the question, and that the two MA interpretations are compositionally derived from LFs with different question nuclei.

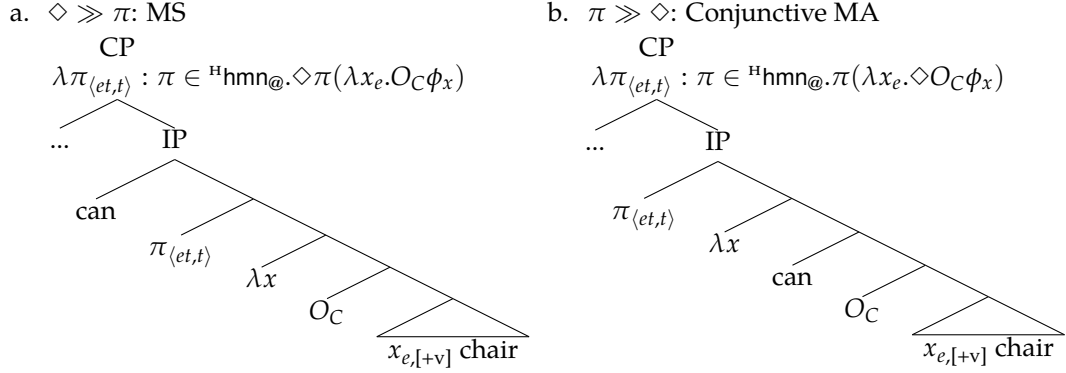
(53) (Only Andy and Billy can chair the committee. Co-chairing is disallowed.)

Who can chair the committee?

- | | |
|-----------------------------|------------------|
| a. Andy can. | (MS) |
| b. Andy can, and Billy can. | (Conjunctive MA) |
| c. Andy or Billy can. | (Disjunctive MA) |

The derivation of the conjunctive MA interpretation is straightforward: it arises if the higher-order *wh*-trace takes scope above the existential modal inside the question nucleus (viz., the IP). Compared with (54a), which is simplified from (50), the only change in (54b) is that the higher-order *wh*-trace π takes scope above *can*.

(54) Who can chair the committee?



The answer spaces yielded by (54a,b) are illustrated in Figures 5a,b, respectively. (Figure 5a is identical to Figure 3.)

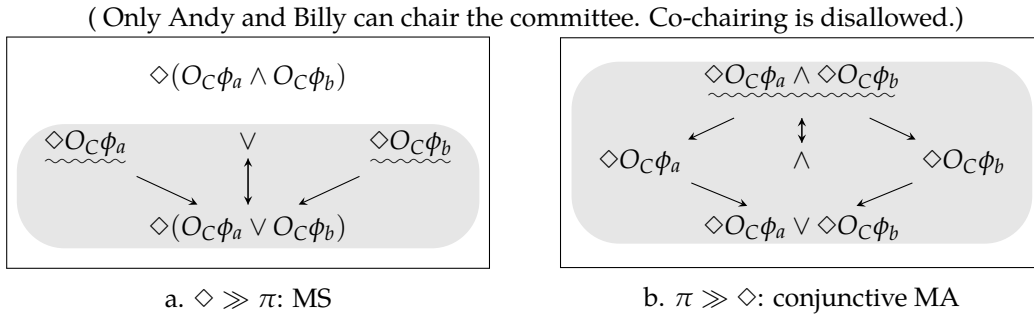


Figure 5: Answer spaces yielded by (54a,b) [Legends and abbreviations are the same as in Figure 3.]

In Figure 5b, the answer space is closed under conjunction: the conjunctive answer is not contradictory and is logically equivalent to the conjunction of the two individual answers. Hence, the resulting interpretation is a MA interpretation. In the described multiple-choice scenario, while the answer space in Figure 5a has multiple max-informative true answers, the one in Figure 5b has only one max-informative true answer, $\diamond O_C\phi_a \wedge \diamond O_C\phi_b$, derived based on $a^\uparrow \cap b^\uparrow$. The other answers are asymmetrically entailed by this conjunctive answer and thus are all partial.

4.4. Deriving disjunctive MA interpretations

MA answers to *can*-questions are more commonly expressed as disjunctions rather than conjunctions. As exemplified in (55), when uttered as a response to a *can*-question, a disjunction may convey either epistemic ignorance or universal free choice (FC), resulting in the use of a partial answer and a mention-few/all answer, respectively.

- (55) Who can teach Intro Chinese?
- a. Andy or Billy (can) ... (but I don't know who). (Ignorance: partial)
 \rightsquigarrow *Either Andy or Billy can teach Intro Chinese, but I don't know who.* ($\diamond\phi_a \vee \diamond\phi_b$)
- b. Andy or Billy (can). (FC: mention-few/all)
 \rightsquigarrow *Andy can teach Intro Chinese, and Billy can teach Intro Chinese, too.* ($\diamond\phi_a \wedge \diamond\phi_b$)

I argue that the ambiguity between MS and disjunctive MA in *can*-questions has the same origin as the ignorance/FC-ambiguity in *can*-disjunctions: in a *can*-question, a disjunctive MA interpretation

arises if and only if felicitous disjunctive answers can be understood as universal FC statements.

There is a rich literature on the derivation of FC in *can*-disjunctions. What this paper demands is an analysis that achieves the following: (i) it derives the FC use of disjunctions while not bringing up additional inferences such as exclusivity, since the disjunctive MA interpretation doesn't force strong exhaustivity; and (ii) it explains why only *can*-questions admit FC-disjunctive answers.²⁰ In the following, I will offer an analysis based on the syntax and semantics of the Mandarin particle *dou*. This particle can function as a FC-trigger in *can*-disjunctions and as an exhaustivity marker in *can*-questions. I will also argue that the modal obviation effect of the FC-trigger use of *dou* and the distributional constraint that only *can*-questions admit universal FC answers have the same origin. My analysis will account for (i) by 'anti-exhaustification' (a la Kratzer and Shimoyama 2002) and explain (ii) based on a definedness condition called 'Relativized Exclusivity'.

4.4.1. Two uses of *dou*: Exhaustivity marker and FC-trigger

The Mandarin particle *dou* has various uses. These uses can be disambiguated by the form of the environment *dou* occurs in and the meaning of the expression associated with *dou*. What concerns us here is that the uses of *dou* in *can*-questions and in *can*-disjunctions exhibit an interesting parallel.

In a *can*-question, as seen in (56a,b), associating *dou* with the *wh*-phrase blocks the MS interpretation. (Underlying marks the expression associated with *dou*.)

- (56) a. (Dou) shui keyi jiao jichu hanyu?
 (DOU) who can teach Intro Chinese
 Without *dou*: 'Who can teach Intro Chinese?' (✓MS, ✓MA)
 With *dou*: 'Who all can teach Intro Chinese?' (✗MS, ✓MA)
- b. Mali (dou) keyi zai nali mai dao kafei?
 Mary (DOU) can at where buy get coffee
 Without *dou*: 'Where can Mary get coffee?' (✓MS, ✓MA)
 With *dou*: 'Where all can Mary get coffee?' (✗MS, ✓MA)

I call *dou* in this use descriptively an 'exhaustivity marker' (after Beck and Rullmann 1999) and assume the LFs in (57) for the above two sentences.²¹ In these LFs, *dou* appears within the IP and is associated with the *wh*-trace in its c-commanding domain.²²

²⁰This paper will be mainly engaged with the (anti-)exhaustification-based analyses of FC (Kratzer and Shimoyama 2002; Chierchia 2006, 2013; Fox 2007; Bar-Lev and Fox 2020; a.o.), which derive FC inferences by applying (anti-)exhaustification. There is also a large set of works that pursue a semantic analysis (Simons 2005; Aloni 2007; Barker 2010; a.o.).

²¹In (56), *dou* is translated as 'all' for convenience. However, *dou* should not be analyzed as a universal distributor (contra Lin 1998, Jie Li 1995, Xiaoguang Li 1997). Xiaoguang Li (1997) assumes that, in its use as an exhaustivity marker, *dou* is associated with a covert adverbial denoting multiple events and quantifies over these events. This analysis cannot explain the unavailability of MS in *can*-questions like (i-a). If *dou* were associated with a covert quantificational adverbial over events here, then (i-a) should have a pair-list/individual MS interpretation, like the ones available in (i-b). However, such interpretation is not available in (i-a); for example, if Starbucks is always accessible to John while Peet's is only sometimes accessible to John, 'Starbucks' is a proper answer to (i-b) but not to (i-a).

- (i) a. Yuehan dou keyi qu nali mai kafei?
 John DOU can go where buy coffee?
 'Where all can John buy coffee?' (MA)
- b. Yuehan mei-ci dou keyi qu nali mai kafei?
 John each-time DOU can go where buy coffee?
 'In each case, where can John can buy coffee?' (Pair-list MS)
 'John always can buy coffee from where?' (Individual MS)

²²The following explains why the exhaustivity marker *dou* has the assumed syntax. First, as seen in (56b) vs. (i), *dou* must

- (57) a. [_{CP} who_i C⁰ [_{IP} dou_C [_{VP} t_{i,[+v]} can teach Intro Chinese]]]
 b. [_{CP} where_i C⁰ [_{IP} Mary_j dou_C [_{VP} t_j can get coffee t_{i,[+v]}]]]

In *can*-declaratives, associating *dou* with a pre-verbal disjunction evokes a universal FC inference, as exemplified in (58). Note that these two sentences do not give rise to an exclusivity inference; for example, even with *dou* present, (58a) doesn't imply that no one else can teach Intro Chinese. This fact argues that FC is independent of exclusivity.

- (58) a. Yuehan huozhe Mali (dou) keyi jiao jichu hanyu.
 John or Mary (DOU) can teach Intro Chinese
 Without *dou*: 'Either John or Mary can teach Intro Chinese.' (Ignorance)
 With *dou*: 'Both John and Mary (as well as possibly others) can teach Intro Chinese.'
 (Universal FC)
- b. Mali zai Xingbake huozhe Maidanglao (dou) keyi mai dao kafei.
 Mary at Starbucks or McDonalds (dou) can buy get coffee
 Without *dou*: 'From either SB or MD, Mary can get coffee.' (Ignorance)
 With *dou*: 'From both SB and MD (as well as possibly other places), Mary can get coffee.'
 (Universal FC)

The FC-trigger use of *dou* is subject to a *modal obviation* effect. As seen in (59), when the existential modal *keyi* 'can' is dropped or replaced with a universal modal like *bixu* 'must', *dou* cannot be grammatically associated with a pre-verbal disjunction. This licensing constraint applies crosslinguistically to universal FC items, including the pre-verbal *any* in English, as seen in (60).

- (59) a. Yuehan huozhe Mali dou keyi/*bixu jiao jichu hanyu.
 John or Mary DOU can/*must teach Intro Chinese
 b. Yuehan huozhe Mali (*dou) jiao -guo jichu hanyu.
 John or Mary (*DOU) teach -EXP Intro Chinese
- (60) a. Anyone can/*must teach Intro Chinese.
 b. * Anyone taught Intro Chinese.

Based on these two uses of *dou*, I argue that a *can*-question has a disjunctive MA interpretation if a *dou*-like operator appears in the question nucleus and is associated with the *wh*-trace. This operator strengthens the disjunctive answers into FC statements.

appear on the right side of the subject unless the subject is an interrogative *wh*-expression. Since a non-interrogative subject stays inside the IP, this fact argues that *dou* is also inside the IP.

- (i) (*Dou) Mali keyi zai nali mai dao kafei?
 (*DOU) Mary can at where buy get coffee

Second, when used as an exhaustivity marker, *dou* must be associated with the *wh*-expression and appear to the left of this *wh*-associate. As seen in (ii), *dou* functions as an exhaustivity marker when appearing before *shenme* 'what' and as a universal distributor when appearing after *shenme*. Given that Mandarin is *wh*-in-situ and assuming that *wh*-items in questions undertake covert *wh*-movement at LF (Huang 1982), I argue that the exhaustivity marker *dou* c-commands the *wh*-trace at LF.

- (ii) (John can give Mary either all of the apples or some (but not all) of the cookies; the choice is up to him.)
- | | |
|--|---|
| a. Yuehan dou keyi ba <u>shenme</u> gei Mali?
John DOU can BA what give Mary
'What all is John allowed to give to Mary?'
Proper reply: 'The apples or some of the cookies.' | b. Yuehan keyi ba <u>shenme</u> dou gei Mali?
John can BA what DOU give Mary
'Which <i>x</i> is s.t. John can give all of <i>x</i> to Mary?'
Proper reply: 'The apples.' |
|--|---|

4.4.2. Deriving the FC-trigger use: *dou* DOU as an anti-exhaustification operator

Kratzer and Shimoyama (2002) derive the FC inference as a result of ‘anti-exhaustification’. Their idea is as follows. The speaker prefers $\diamond(\phi \vee \psi)$ to the two stronger alternatives $\diamond\phi$ and $\diamond\psi$ because she is unhappy with the strengthened meanings of these alternatives — the reason cannot be that the speaker is unhappy with the non-strengthened meanings of these alternatives, because negating both $\diamond\phi$ and $\diamond\psi$ would yield an inference that contradicts $\diamond(\phi \vee \psi)$. Hence, uttering $\diamond(\phi \vee \psi)$ implicates that the strengthened meaning of each stronger alternative is false, which is why the proposed operation is called ‘anti-exhaustification’.

Xiang (2020) analyzes the Mandarin particle *dou* as an anti-exhaustification operator. Roughly, as schematized in (61), *dou* affirms the prejacent and negates the (innocent-exclusion (IE-)based) exhaustification of each ‘non-innocently-excludable alternative’ except the prejacent itself.^{23,24}

$$(61) \text{ The asserted meaning of } dou \text{ (simplified from Xiang 2020; presuppositions are ignored)} \\ \llbracket dou_C \rrbracket = \lambda p \lambda w : p(w) = 1 \wedge \underbrace{\forall q \in (C - \{p\}) [q \notin \text{IE}_{\text{EXCL}}(p, C) \rightarrow O_C^{\text{IE}}(q)(w) = 0]}_{\text{anti-exhaustification}}$$

As defined by Fox (2007), an alternative is ‘innocently (I-)excludable’ if and only if it is included in every maximal set of alternatives A such that affirming the prejacent is consistent with negating all the alternatives in A . ‘Innocent exclusion (IE-)based exhaustification’ is the exhaustification operation that negates only the I-excludable alternatives (cf. the exhaustification O -operator, which negates all the non-entailed alternatives, defined in (46)).

$$(62) \text{ a. Innocently (I-)excludable alternatives (Fox 2007)} \\ \text{IE}_{\text{EXCL}}(p, C) = \bigcap \{A \mid A \text{ is a maximal subset of } C \text{ s.t. } \{\neg q \mid q \in A\} \cup \{p\} \text{ is consistent}\} \\ \text{b. Innocent exclusion (IE-)based exhaustification} \\ O_C^{\text{IE}} = \lambda q \lambda w. q(w) = 1 \wedge \forall r \in \text{IE}_{\text{EXCL}}(q, C) [r(w) = 0]$$

Innocent exclusion differs from traditional exclusion mainly in sentences with disjunctions or existential quantifiers. In a disjunctive sentence, the connective *or* carries a [+D] feature, which activates a set of domain alternatives, defined as follows:

$$(63) \text{ D-ALT}(or_{[+D]}) = \{\lambda b \lambda a. a \sqcup b, \lambda b \lambda a. a, \lambda b \lambda a. b\}$$

For the narrow-scope \diamond -disjunction $\diamond(\phi \vee \psi)$, the subdomain alternatives $\diamond\phi$ and $\diamond\psi$ are not I-excludable because $\{\neg\diamond\phi, \neg\diamond\psi\} \cup \{\diamond(\phi \vee \psi)\}$ is inconsistent (or say, $\diamond(\phi \vee \psi) \wedge \neg\diamond\phi \wedge \neg\diamond\psi = \perp$).

²³In addition to the asserted meaning in (61), Xiang 2020 also assumes a non-vacuity presupposition, namely that the sentential argument of *dou* has at least one alternative that can participate in anti-exhaustification. This presupposition is omitted in this paper since it doesn’t affect the FC-trigger use or the exhaustivity marker use of *dou*.

²⁴What is particularly advantageous about defining *dou* as an anti-exhaustification operator is that it accounts for not only the FC-trigger use of *dou* but also a few other uses. In Xiang 2020, the alternatives that participate in anti-exhaustification are uniformly called ‘sub-alternatives’. The function of *dou* varies depending on what alternatives count as sub-alternatives. Briefly, sub-alternatives are primarily alternatives that are logically weaker than the prejacent, but in certain syntactic and prosodic environments, the ordering source in defining sub-alternatives gets shifted from logical strength to innocent excludability or likelihood, yielding different functions of *dou*. This analysis accounts for the universal distributor use, the FC-trigger use, as well as the *even*-like scalar marker use. In particular, the FC-trigger use of *dou* arises when sub-alternatives are defined in terms of innocent excludability, per the definition in (61).

Mingming Liu (pers. comm.) argues that the Mandarin particle *dou* itself doesn’t assert anti-exhaustivity; instead, he defines *dou* as an *even*-like operator (Liu 2016a) and argues that the presupposition of *dou/even* — that the prejacent has a more likely alternative — forces a separate anti-exhaustification operator to apply in order to strengthen the prejacent disjunction into a FC statement. This idea is reviewed in Xiang 2020: Appendix B. For present concerns, it only matters that associating *dou* with a disjunction forces anti-exhaustification.

For an alternative account that derives the FC-trigger use of *dou* without assuming anti-exhaustification, see Zhao 2019.

In contrast, for the narrow-scope \Box -disjunction $\Box(\phi \vee \psi)$, the subdomain alternatives $\Box\phi$ and $\Box\psi$ are I-excludable because $\Box(\phi \vee \psi) \wedge \neg\Box\phi \wedge \neg\Box\psi \neq \perp$. A summary is given in Table 2.

	Anti-excludable		I-excludable
	Weaker	Neither	
$\phi \wedge \psi$ $\Box\phi \wedge \Box\psi, \Box(\phi \wedge \psi)$ $\Diamond\phi \wedge \Diamond\psi, \Diamond(\phi \wedge \psi)$	ϕ, ψ $\Box\phi, \Box\psi$ $\Diamond\phi, \Diamond\psi$		
$\phi \vee \psi$ $\Diamond\phi \vee \Diamond\psi, \Diamond(\phi \vee \psi)$ $\Box\phi \vee \Box\psi$ $\Box(\phi \vee \psi)$		ϕ, ψ $\Diamond\phi, \Diamond\psi$ $\Box\phi, \Box\psi$	$\Box\phi, \Box\psi$

Table 2: Subdomain alternatives of conjunctions and disjunctions [‘Neither’ means ‘neither weaker nor I-excludable’.]

I henceforth call the alternatives that participate in anti-exhaustification ‘anti-excludable alternatives’ (abbreviation: ANTIEXCL).²⁵ The semantics in (61) is re-written as follows:

$$(64) \quad \llbracket dou_C \rrbracket / \llbracket DOU_C \rrbracket = \lambda p \lambda w : p(w) = 1 \wedge \forall q \in \text{ANTIEXCL}(p, C) [O_C^E(q)(w) = 0]$$

where $\text{ANTIEXCL}(p, C) = (C - \text{IEXCL}(p, C)) - \{p\}$

Note that weaker alternatives are anti-excludable, although the inference yielded by anti-exhaustifying the weaker alternatives is entailed by the prejacent and thus doesn’t affect the truth conditions. According to the definition in (64), the application of *dou* is semantically vacuous unless the prejacent sentence has alternatives that are neither weaker nor I-excludable, labeled in Table 2 as ‘Neither’.

The FC inference of (58a) is now computed as in (65). Associated with a disjunction, *dou* quantifies over the set of domain alternatives of its prejacent, which includes two anti-excludable alternatives (viz., the two modalized disjuncts). Employing *dou* affirms the prejacent disjunctive sentence and negates the exhaustification of each disjunct, yielding a conjunctive/universal FC inference.

$$(65) \quad [dou_C [s \text{ John or}_{[+D]} \text{ Mary can teach Intro Chinese}]]$$

- a. $\llbracket S \rrbracket = \Diamond\phi_j \vee \Diamond\phi_m$ (ϕ_x abbreviates ‘ x teaches Intro Chinese’)
- b. $C = \text{D-ALT}(S) = \{\Diamond\phi_j, \Diamond\phi_m, \Diamond\phi_j \vee \Diamond\phi_m\}$
- c. $\text{ANTIEXCL}(\llbracket S \rrbracket, C) = \{\Diamond\phi_j, \Diamond\phi_m\}$
- d. $\llbracket dou_C(S) \rrbracket \Leftrightarrow [\Diamond\phi_j \vee \Diamond\phi_m] \wedge \neg O_C^E \Diamond\phi_j \wedge \neg O_C^E \Diamond\phi_m$
 $\Leftrightarrow [\Diamond\phi_j \vee \Diamond\phi_m] \wedge [\Diamond\phi_j \rightarrow \Diamond\phi_m] \wedge [\Diamond\phi_m \rightarrow \Diamond\phi_j]$
 $\Leftrightarrow [\Diamond\phi_j \vee \Diamond\phi_m] \wedge [\Diamond\phi_j \leftrightarrow \Diamond\phi_m]$
 $\Leftrightarrow \Diamond\phi_j \wedge \Diamond\phi_m$

Due to separate syntactic constraints in Mandarin, *dou* cannot be associated with a narrow-scope disjunction in overt syntax. For example, it cannot grammatically occur in the Mandarin translation of ‘I (**dou*) can ask John or Mary to teach Intro Chinese’ (*can* \gg *or*). Yet nothing prevents *dou* from being associated with a narrow-scope expression at LF. In light of Matthewson’s (2001) ‘No Variation Hypothesis’ that semantics has no crosslinguistic variation, I assume that anti-exhaustification is

²⁵I don’t call such alternatives ‘innocently (I-)includable alternatives’, a notion coined by Bar-Lev and Fox (2020), because I-includable alternatives include also the prejacent itself (for details, see (92) in Sect. 5.2.1).

a crosslinguistically available source for universal FC. I also assume that anti-exhaustification is realized by the application of either overt *dou* or covert DOU .

It's worth discussing how the presented analysis of FC differs from the analyses of Fox 2007 and Chierchia 2006, 2013, which also adapt anti-exhaustification to the grammatical approach of implicatures. Fox (2007) analyzes anti-exhaustivity as a consequence of applying IE-based exhaustification recursively. This analysis predicts that FC and exclusivity arise at the same time (the same goes for Bar-Lev and Fox 2020, see Sect. 5.2.1). In contrast, in light of the data in (58), I derive these two inferences separately by treating exhaustification and anti-exhaustification as two independent operations: FC stems from the application of the anti-exhaustification operator DOU , while exclusivity stems from the application of a separate IE-based exhaustification operator. This difference also affects question semantics: applying recursive exhaustification to the question nucleus inevitably yields strong exhaustivity, while applying DOU allows for weak exhaustivity.

Chierchia (2006, 2013) considers the choice between exhaustification and anti-exhaustification a result of syntactic agreement: whether the alternatives participate in exhaustification or anti-exhaustification depends on whether the alternative-triggering feature (e.g., [+D]) agrees with a basic exhaustifier O_D or a pre-exhaustification exhaustifier $O_{D\text{-EXH}}$. In Chierchia's analysis, the subdomain alternatives of $\Box(\phi \vee \psi)$, namely $\Box\phi$ and $\Box\psi$, participate in anti-exhaustification if and only if the [+D] feature of the disjunction agrees with a pre-exhaustification exhaustifier in syntax. In contrast, following Kratzer and Shimoyama (2002) and Fox (2007), I assume that whether an alternative may participate in anti-exhaustification is purely determined by its semantics, namely, whether it is non-I-excludable with respect to the prejacent. For example, the application of DOU to $\Box(\phi \vee \psi)$ is semantically vacuous since $\Box\phi$ and $\Box\psi$ are I-excludable with respect to $\Box(\phi \vee \psi)$.

4.4.3. Modal obviation of FC: Relativized Exclusivity

If anti-exhaustification had no application constraint, it would derive universal FC for a variety of disjunctive sentences, such as the non-modalized disjunction $\phi \vee \psi$ and any wide-scope disjunctions, including $\Box\phi \vee \Box\psi$. For example, affirming $\phi \vee \psi$ and anti-exhaustifying the two alternatives ϕ and ψ would return $\phi \wedge \psi$. This prediction is obviously wrong. Hence, there must be constraints on when anti-exhaustification is available — constraints which result in the modal obviation effect that universal FC is only possible in sentences with an existential modal (i.e., $\Diamond(\phi \vee \psi)$ and $\Diamond\phi \vee \Diamond\psi$).

In the realm of exhaustification-based theories of FC, explanations of the modal obviation effect fall into two groups. One attributes the unavailability of FC in $\phi \vee \psi$ to the contradiction between the FC inference $\phi \wedge \psi$ and the scalar implicature $\neg(\phi \wedge \psi)$ (Chierchia 2013; Bar-Lev and Fox 2020). The other relates the (in-)compatibility of interpreting the subdomain alternatives to local exhaustification (Menéndez-Benito 2010; Dayal 2013; Xiang 2020): for example, $O\phi$ and $O\psi$ are mutually exclusive, while $\Diamond O\phi$ and $\Diamond O\psi$ are not. This paper doesn't pursue the first strategy because it doesn't account for the distribution of FC-disjunctive answers in singular *wh*-questions. (For details, see the review of Fox's (2018, 2020) analysis in Sect. 5.2.1, which uses this strategy.)

Following the second strategy, I propose that the anti-exhaustification operator *dou*/ DOU has a 'Relativized Exclusivity (*RelExcl*)' presupposition (cf. the *Viability* constraint of Dayal 2013): for every anti-excludable alternative ϕ stronger than the prejacent, every minimal set of accessible worlds that verifies ϕ also verifies the (IE-based) exhaustification of ϕ .^{26,27} This presupposition is schematized as

²⁶In this paper, *RelExcl* is treated as a definedness condition of anti-exhaustification. However, the effects of *RelExcl* are independent of the assumptions about anti-exhaustification; the only prerequisite for assuming *RelExcl* is that the formal theory used for deriving universal FC allows us to extract the subdomain alternatives from a disjunctive sentence.

²⁷For a set of possible worlds W and a sentence ϕ , we have:

follows, where M is a modal base, and C/C' denotes a subset of the alternatives of $\llbracket S \rrbracket^M / \llbracket S \rrbracket^{M'}$:^{28,29}

(66) *Relativized Exclusivity*

$\llbracket \text{dou}_C(S) \rrbracket^M / \llbracket \text{DOU}_C(S) \rrbracket^M$ is defined in w only if

$\forall \phi [\llbracket \phi \rrbracket^M \in \text{ANTIEXCL}(\llbracket S \rrbracket^M, C) \wedge \llbracket \phi \rrbracket^M \subset \llbracket S \rrbracket^M$

$\rightarrow \forall M'_{\langle s, st \rangle} [M'_w$ is a minimal subset of M_w s.t. $\llbracket \phi \rrbracket^{M'}(w) = 1 \rightarrow O_{C'}^{\text{IE}}(\llbracket \phi \rrbracket^{M'})(w) = 1]$

(For any sentence ϕ such that $\llbracket \phi \rrbracket^M$ is an anti-excludable alternative of S stronger than $\llbracket S \rrbracket^M$, we have: for every modal base M' such that M'_w is a minimal set of w -accessible worlds that verifies ϕ in w , M'_w also verifies the exhaustification of ϕ in w .)

The rest of this subsection will explain how the RelExcl presupposition accounts for modal obviation. For simplicity, the formula $O_{C'}^{\text{IE}}\phi$ (viz., the IE-based exhaustification of ϕ relative to the alternative set C) will be abbreviated as $O\phi$.

For a non-modalized disjunction, the choice of modal base makes no difference. For (67), RelExcl simply means that $\phi_j \wedge \phi_m$, $O\phi_j$, and $O\phi_m$ are simultaneously true, which is clearly contradictory.

(67) * DOU_C [John or Mary teach Intro Chinese]

(Ungrammatical because $\text{DOU}_C[\phi_j \vee \phi_m]$ inevitably violates RelExcl.)

For a wide-scope \square -disjunction, there is no modal base that allows the universal FC inference and RelExcl presupposition to be simultaneously true. For (68), we have: (i) the FC inference $\square\phi_j \wedge \square\phi_m$ is true in w relative to M iff John and Mary both teach Intro Chinese in every world in M_w ; (ii) $O\square\phi_j$ is true in w relative to M' iff only John teaches Intro Chinese in every world in M'_w ; (iii) clearly M'_w cannot be a subset of M_w unless $M'_w = \emptyset$.

(68) * DOU_C [John or Mary must teach Intro Chinese]

(Ungrammatical because $\text{DOU}_C[\square\phi_j \vee \square\phi_m]$ inevitably violates RelExcl.)

- (i) a. W verifies $\diamond\phi$ in w if and only if there is a w -accessible world w' in W such that ϕ is true in w' ;
b. W verifies $\square\phi$ in w if and only if for every w -accessible world w' in W , ϕ is true in w' .

For illustration, consider three w -accessible worlds w_1, w_2, w_3 , and assume that the sets of individuals who teach Intro Chinese in these three worlds are $\{j\}$, $\{m\}$, and $\{j, m\}$, respectively. Let ϕ_x abbreviate ' x teaches Intro Chinese'; then the sentences and the minimal sets of accessible worlds that verify these sentences are paired as follows:

- (ii) a. $\diamond\phi_j, \square\phi_j: \{w_1\}, \{w_3\}$
b. $\diamond\phi_m, \square\phi_m: \{w_2\}, \{w_3\}$
c. $\diamond\phi_j \wedge \diamond\phi_m: \{w_1, w_2\}, \{w_3\}$
d. $\diamond\phi_j \vee \diamond\phi_m, \diamond(\phi_j \vee \phi_m), \square\phi_j \vee \square\phi_m, \square(\phi_j \vee \phi_m): \{w_1\}, \{w_2\}, \{w_3\}$
e. $\square\phi_j \wedge \square\phi_m, \square(\phi_j \wedge \phi_m): \{w_3\}$

²⁸Note that here ϕ stands for a syntactic expression, not a semantic value. The RelExcl condition is formalized as such because it concerns the interpretations of the alternatives evaluated relative to different modal bases. However, this definition is syncategorematic: it cannot be written as a lexical constraint (viz., cannot be written as ' $\llbracket \text{DOU}_C \rrbracket = \dots$ '), and it cannot give a precise characterization of C' . To solve these problems, I will explore a variable-free treatment of modal bases in Appendix B.

²⁹The RelExcl condition is very similar to Dayal's (2013) *Viability* constraint, which says that every exhaustified alternative is true relative to a subset of the accessible worlds. The following formulates this constraint analogously to (66):

- (i) $\forall \phi [\llbracket \phi \rrbracket^M \in \text{ANTIEXCL}(\llbracket S \rrbracket^M, C) \wedge \llbracket \phi \rrbracket^M \subset \llbracket S \rrbracket^M \rightarrow \exists M' [M'_w \subseteq M_w \wedge O_{C'}^{\text{IE}}(\llbracket \phi \rrbracket^{M'})(w) = 1]$

(For any sentence ϕ such that $\llbracket \phi \rrbracket^M$ is an anti-excludable alternative of S stronger than $\llbracket S \rrbracket^M$, there is a modal base M' such that M'_w is a subset of w -accessible worlds that verifies the exhaustification of ϕ in w .)

The main difference between RelExcl and Viability is the following: in (69b), where the \diamond -disjunction is parsed without local exhaustification, Viability predicts that universal FC is possible as long as $\diamond O\phi_j$ and $\diamond O\phi_m$ are true, regardless of the truth or falsity of $\diamond(\phi_j \wedge \phi_m)$. However, as we shall see in Sect. 6.3.4, to account for the universal local-uniqueness inferences in *can*-questions with a disjunctive MA interpretation, there has to be an interpretation that requires $\diamond(\phi_j \wedge \phi_m)$ to be false.

As for the corresponding narrow-scope \Box -disjunction, the application of DOU is vacuous: for $\Box(\phi_j \vee \phi_m)$, the alternatives $\Box\phi_j$ and $\Box\phi_m$ are I-excludable and thus do not participate in anti-exhaustification.

The presence of an existential modal salvages the violation of RelExcl. Example (69) demonstrates two ways to parse a wide-scope \Diamond -disjunction, namely, without and with local exhaustification. In these two parses, RelExcl yields two definedness conditions which differ w.r.t. whether there is an accessible world where both John and Mary teach Intro Chinese. As stated in (69), assume that only John, only Mary, and they both teach Intro Chinese in w_1 , w_2 , and w_3 , respectively. Given the assumed modal base M , the FC inference $\Diamond\phi_j \wedge \Diamond\phi_m$ is true in all three anchor worlds w, w', w'' .

(69) DOU_C [John or Mary can teach Intro Chinese]

$$\text{Let teach-IC} = \begin{bmatrix} w_1 \rightarrow \{j\} \\ w_2 \rightarrow \{m\} \\ w_3 \rightarrow \{j, m\} \end{bmatrix} \text{ and } M = \begin{bmatrix} w \rightarrow \{w_1, w_2\} \\ w' \rightarrow \{w_1, w_2, w_3\} \\ w'' \rightarrow \{w_1, w_3\} \end{bmatrix}$$

- a. Given M , RelExcl is satisfied in w but violated in w' and w'' . More generally:
 $\text{DOU}_C[\Diamond\phi_j \vee \Diamond\phi_m] = \Diamond\phi_j \wedge \Diamond\phi_m$, defined only if $\Diamond O\phi_j \wedge \Diamond O\phi_m \wedge \neg\Diamond(\phi_j \wedge \phi_m)$.
- b. Given M , RelExcl is satisfied in w and w' but violated in w'' . More generally:
 $\text{DOU}_C[\Diamond O\phi_j \vee \Diamond O\phi_m] = \Diamond O\phi_j \wedge \Diamond O\phi_m$, defined only if $\Diamond O\phi_j \wedge \Diamond O\phi_m$.

If the sentence is parsed without local exhaustification, as in (69a), RelExcl is satisfied only in w . The minimal set of w -accessible worlds that verifies $\Diamond\phi_j$, namely $\{w_1\}$, also verifies its exhaustification $O\Diamond\phi_j$; likewise for $\Diamond\phi_m$, the set $\{w_2\}$, which verifies $\Diamond\phi_m$, also verifies the exhaustification $O\Diamond\phi_m$. In contrast, RelExcl is not satisfied in w' , which has one more accessible world w_3 : $\{w_3\}$ verifies $\Diamond\phi_j$ and $\Diamond\phi_m$ but not $O\Diamond\phi_j$ or $O\Diamond\phi_m$.

However, if the sentence is parsed with local exhaustification, as in (69b), RelExcl is also satisfied in w' : the violation of exclusivity in w_3 does not affect RelExcl because $\{w_3\}$ does not verify any of the locally exhaustified disjuncts (viz., $\Diamond O\phi_j$ or $\Diamond O\phi_m$). In this case, the definedness condition yielded by RelExcl is equivalent to the FC inference itself.

This analysis also applies to the narrow-scope \Diamond -disjunction:

- (70) a. $\text{DOU}_C[\Diamond(\phi_j \vee \phi_m)] = \Diamond\phi_j \wedge \Diamond\phi_m$, defined only if $\Diamond O\phi_j \wedge \Diamond O\phi_m \wedge \neg\Diamond(\phi_j \wedge \phi_m)$.
- b. $\text{DOU}_C[\Diamond(O\phi_j \vee O\phi_m)] = \Diamond O\phi_j \wedge \Diamond O\phi_m$, defined only if $\Diamond O\phi_j \wedge \Diamond O\phi_m$.

To sum up, anti-exhaustification is a crosslinguistic source for universal FC. This operation is lexicalized as *dou*/ DOU . When associated with a disjunction, *dou*/ DOU affirms the prejacent and negates the IE-based exhaustification of each anti-excludable alternative, giving rise to a universal FC inference. Moreover, the anti-exhaustification operator *dou*/ DOU presupposes Relativized Exclusivity (RelExcl). In a disjunctive sentence, the RelExcl presupposition can be satisfied only in the presence of an existential modal, which may appear either above or below the disjunction.

4.4.4. Deriving disjunctive MA

Drawing on the parallel uses of *dou* in disjunctions and *wh*-questions, I argue that the disjunctive MA interpretations of *can*-questions in English are derived by applying a covert anti-exhaustification operator DOU within the question nucleus. When used in *wh*-questions, this DOU -operator has the same LF syntax as the Mandarin exhaustivity marker *dou*, namely, it appears within the question nucleus and is associated with a *c*-commanding *wh*-trace (argued in fn. 22).

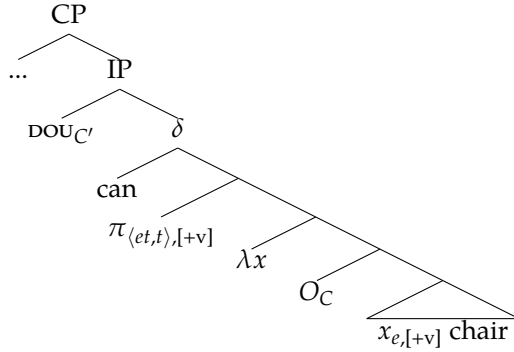
The resulting LFs are structured as in (71). The δ -nodes in (71a) and (71b) are simply the IP nodes

in the LF of higher-order MS and the LF of conjunctive MA, respectively. These two δ -nodes only differ in the scope of the higher-order *wh*-trace π relative to the modal verb *can*. In both LFs, a covert anti-exhaustification operator DOU is applied to the δ -node and is associated with the *wh*-trace π . The variable C' carried by DOU denotes a set of variable alternatives of δ , which is equivalent to the answer space derived in the absence of DOU .

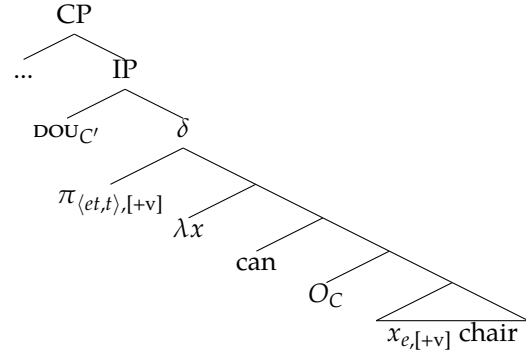
(71) 'Who can chair the committee?' (Disjunctive MA)

'Andy or Billy can.'
 $\sim \sim$ *Andy can chair the committee, and Billy can chair the committee.*

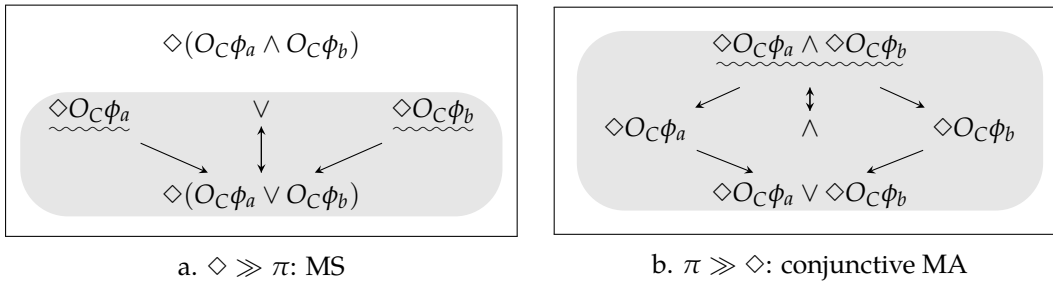
a. $\text{DOU} \gg \diamond \gg \pi$



b. $\text{DOU} \gg \pi \gg \diamond$



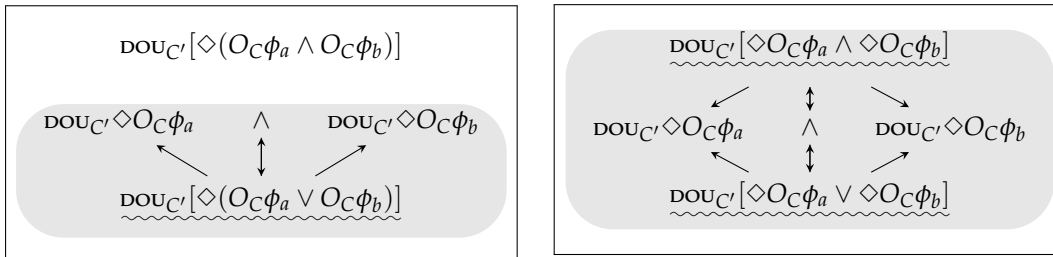
Compare now the answer spaces yielded in the absence of DOU , which are displayed once more in Figures 6a,b, to the answer spaces yielded by the LFs in (71a,b), depicted in Figures 7a,b. Legends, abbreviations, and the setup of the context are the same as in Figure 3.³⁰



a. $\diamond \gg \pi$: MS

b. $\pi \gg \diamond$: conjunctive MA

Figure 6: Answer spaces of (71a,b) yielded in the absence of DOU



a. $\text{DOU} \gg \diamond \gg \pi$: disjunctive MA

b. $\text{DOU} \gg \pi \gg \diamond$: con/dis-junctive MA

Figure 7: Answer spaces of (71a,b) yielded in the presence of DOU

³⁰Figure 7b is labeled as 'dis/con-junctive MA', because the conjunctive answer at the top and the disjunctive answer at the bottom are truth-conditionally equivalent. They both express the MA answer that Andy alone and Billy alone can chair.

The computation of the alternatives in Figures 7a,b are given in (72) and (73) below. As defined in (72a) and (73a), the variables C and C' are the domain variables carried by the local exhaustification operator O and the global anti-exhaustification operator DOU , respectively. In both cases, the applications of DOU in the individual answers and the conjunctive answers are semantically vacuous: the individual answers have no anti-excludable alternative; the conjunctive answers have anti-excludable alternatives, but since these alternatives are weaker than the prejacent, the anti-exhaustification inference (underlined in (72b) and (73b)) is entailed by the prejacent. However, as computed in (72c) and (73c), in both cases, DOU strengthens the disjunctive answer into a universal FC statement that is logically equivalent to the conjunction of the two individual answers, making the answer space closed under conjunction.

(72) For $\text{DOU} \gg \diamond \gg \pi$ (answer space in Figure 7a):

a. $C = \{\phi_x \mid x \in \text{hmn}_@ \}, C' = \{\diamond \pi(\lambda x_e. O_C \phi_x) \mid \pi \in {}^{\text{H}}\text{hmn}_@ \}$

b. Computation of the conjunctive answer

i. $\text{ANTIEXCL}(\diamond(O_C \phi_a \wedge O_C \phi_b), C') = C' - \{\perp\}$

ii. $\text{DOU}_{C'}[\diamond(O_C \phi_a \wedge O_C \phi_b)]$

(Intended: '#It's possible that we have only A and only B to chair the committee.')

$$\Leftrightarrow \diamond(O_C \phi_a \wedge O_C \phi_b) \wedge \forall \phi \in (C' - \{\perp\}) [\text{---}O_C^{\text{IE}} \phi]$$

$$\Leftrightarrow \diamond(O_C \phi_a \wedge O_C \phi_b)$$

$$\Leftrightarrow \perp$$

c. Computation of the disjunctive answer

i. $\text{ANTIEXCL}(\diamond(O_C \phi_a \vee O_C \phi_b), C') = \{\diamond O_C \phi_a, \diamond O_C \phi_b\}$

ii. $\text{DOU}_{C'}[\diamond(O_C \phi_a \vee O_C \phi_b)]$

(Intended: 'We can have A alone or B alone to chair the committee.' (*can* \gg *or*))

$$\Leftrightarrow \diamond(O_C \phi_a \vee O_C \phi_b) \wedge \text{---}O_C^{\text{IE}} \diamond O_C \phi_a \wedge \text{---}O_C^{\text{IE}} \diamond O_C \phi_b$$

$$\Leftrightarrow \diamond(O_C \phi_a \vee O_C \phi_b) \wedge [\diamond O_C \phi_a \rightarrow \diamond O_C \phi_b] \wedge [\diamond O_C \phi_b \rightarrow \diamond O_C \phi_a]$$

$$\Leftrightarrow \diamond(O_C \phi_a \vee O_C \phi_b) \wedge [\diamond O_C \phi_a \leftrightarrow \diamond O_C \phi_b]$$

$$\Leftrightarrow \diamond O_C \phi_a \wedge \diamond O_C \phi_b$$

(73) For $\text{DOU} \gg \pi \gg \diamond$ (answer space in Figure 7b):

a. $C = \{\phi_x \mid x \in \text{hmn}_@ \}, C' = \{\pi(\lambda x_e. \diamond O_C \phi_x) \mid \pi \in {}^{\text{H}}\text{hmn}_@ \}$

b. Computation of the conjunctive answer

i. $\text{ANTIEXCL}(\diamond O_C \phi_a \wedge \diamond O_C \phi_b, C') = \{\diamond O_C \phi_a, \diamond O_C \phi_b\}$

ii. $\text{DOU}_{C'}[\diamond O_C \phi_a \wedge \diamond O_C \phi_b]$

(Read as: 'A alone and B alone can chair the committee.')

$$\Leftrightarrow [\diamond O_C \phi_a \wedge \diamond O_C \phi_b] \wedge \text{---}O_C^{\text{IE}} \diamond O_C \phi_a \wedge \text{---}O_C^{\text{IE}} \diamond O_C \phi_b$$

$$\Leftrightarrow \diamond O_C \phi_a \wedge \diamond O_C \phi_b$$

c. Computation of the disjunctive answer

i. $\text{ANTIEXCL}(\diamond O_C \phi_a \vee \diamond O_C \phi_b, C') = \{\diamond O_C \phi_a, \diamond O_C \phi_b\}$

ii. $\text{DOU}_{C'}[\diamond O_C \phi_a \vee \diamond O_C \phi_b]$

(Read as: 'A alone or B alone DOU can chair the committee.' (*or* \gg *can*))

$$\Leftrightarrow [\diamond O_C \phi_a \vee \diamond O_C \phi_b] \wedge \text{---}O_C^{\text{IE}} \diamond O_C \phi_a \wedge \text{---}O_C^{\text{IE}} \diamond O_C \phi_b$$

$$\Leftrightarrow [\diamond O_C \phi_a \vee \diamond O_C \phi_b] \wedge [\diamond O_C \phi_a \rightarrow \diamond O_C \phi_b] \wedge [\diamond O_C \phi_b \rightarrow \diamond O_C \phi_a]$$

$$\Leftrightarrow [\diamond O_C \phi_a \vee \diamond O_C \phi_b] \wedge [\diamond O_C \phi_a \leftrightarrow \diamond O_C \phi_b]$$

$$\Leftrightarrow \diamond O_C \phi_a \wedge \diamond O_C \phi_b$$

This analysis straightforwardly explains why only *can*-questions admit universal FC-disjunctive answers. In the above *can*-question, the RelExcl presupposition of *dou* yields a definedness condition equivalent to the FC meaning of the disjunctive answer. However, disjunctive answers in other types of *wh*-questions cannot be strengthened into FC statements via anti-exhaustification. For example, for the \square -question in (41), repeated below, a disjunctive answer may express either ignorance ($\square\phi_f \vee \square\phi_r$) or existential FC ($\square[\phi_f \vee \phi_r]$), but not universal FC ($\square\phi_f \wedge \square\phi_r$): applying *dou* to the disjunctive answer either causes an inevitable violation of RelExcl or is semantically vacuous.

- (74) Q: ‘What does John have to read?’ A: ‘The French novels or the Russian novels.’
- a. If $\pi \gg \square$: ignorance
 $\text{DOU}_C[\square\phi_f \vee \square\phi_r]$ isn’t a possible answer, because it inevitably violates RelExcl.
 - b. If $\square \gg \pi$: existential FC
 - i. $\text{DOU}_C[\square(\phi_f \vee \phi_r)] = \square(\phi_f \vee \phi_r)$; anti-exhaustification is vacuous because $\square(\phi_f \vee \phi_r)$ has no anti-excludable alternative.
 - ii. The domain alternatives have to participate in exhaustification, which yields existential FC: $O_C^{\text{E}}[\square(\phi_f \vee \phi_r)] = \square(\phi_f \vee \phi_r) \wedge \neg\square\phi_f \wedge \neg\square\phi_r = \square(\phi_f \vee \phi_r) \wedge \diamond\phi_f \wedge \diamond\phi_r$

4.5. Interim summary

This section has presented a nucleus-dependent approach to deriving the MS and MA interpretations of *can*-questions. I have argued that we should attribute the MS/MA ambiguity in *can*-questions to minimal structural variations within the question nucleus. By adopting the definition of answerhood from Fox 2013, which allows for multiple max-informative true answers, the presented approach predicts that a *can*-question has a MS interpretation unless one of the following conditions is met: (i) the higher-order *wh*-trace takes scope above *can*, or (ii) an anti-exhaustification operator (\approx the Mandarin particle *dou*) appears above *can* and is associated with the higher-order *wh*-trace. In particular, condition (i) yields conjunctive MA, and condition (ii) yields disjunctive MA.

Compared with earlier nucleus-dependent analyses (George 2011: Chap. 6; Fox 2013), this approach has the following merits. First, it accounts for several linguistic properties of MS answers, namely local exhaustivity, mutual independence, and ‘mention-one-only’. Second, by assuming higher-order *wh*-quantification, this approach derives a scopal effect in *can*-questions for free; this scopal effect naturally accounts for the contrast between MS and conjunctive MA. Third, drawing on evidence from the Mandarin multi-functional particle *dou*, this account uniformly explains the derivation and distribution of universal FC disjunctions and disjunctive MA interpretations.

However, as admitted in Sect. 4.2.3, the assumptions made in this section only partially explain the construction-specific MS-licensing effect of *can* — they don’t fully rule out MS interpretations for non-modalized questions and questions with a universal modal. As I will argue in the next two sections, this MS over-generation problem, as well as the *prima facie* dilemma between uniqueness and MS, can all be resolved by ‘Relativized Exhaustivity’.

5. The dilemma

My analysis of the MS/MA ambiguity adopts the definition of answerhood from Fox (2013), which doesn’t demand global exhaustivity. This definition of answerhood, however, conflicts with Dayal’s (1996) exhaustivity presupposition, which is crucial in accounting for uniqueness effects in questions. Moreover, without further constraints, this definition over-predicts MS interpretations for a variety

kinds of questions.

Hirsch and Schwarz (2020) observe that the uniqueness inference triggered by a singular *which*-phrase can be interpreted locally. This observation challenges Dayal’s exhaustivity presupposition as well as a modified exhaustivity presupposition offered by Fox (2018, 2020). They argue to abandon Dayal’s exhaustivity presupposition and propose that the observed uniqueness effects are due to the lexical features of the determiner *which*. This section discusses the resulting dilemma and reviews the two alternative accounts of uniqueness by Fox (2018, 2020) and Hirsch and Schwarz (2020).

5.1. Dayal’s exhaustivity presupposition: merits and challenges

5.1.1. Uniqueness and singular *wh*-questions

Wh-questions with a singular *which*-phrase (called ‘singular *wh*-questions’) are subject to a uniqueness effect. For example, question (75) can be felicitously uttered only if the speaker believes that this question has just one true answer.

(75) Which child came? \rightsquigarrow Only one of the children came.

This uniqueness effect is standardly explained by Dayal’s exhaustivity presupposition, which says that a question is defined only if it has an exhaustive true answer. Dayal (1996) encodes this exhaustivity requirement of question semantics as a presupposition of the answerhood operator, defined as follows:

$$(76) \text{Ans}_{\text{Dayal}}(w)(Q) = \exists p[w \in p \in Q \wedge \forall q[w \in q \in Q \rightarrow p \subseteq q]]. \quad (= (28))$$

$$i p[w \in p \in Q \wedge \forall q[w \in q \in Q \rightarrow p \subseteq q]]$$

She argues that the contrast in uniqueness between singular and plural *wh*-questions comes from the semantic distinction between singular and plural nouns: a singular noun denotes a set of atomic entities, while a plural noun denotes a set that also includes sums (Sharvy 1980; Link 1983). Incorporating this distinction into *wh*-questions, Dayal (1996) argues that the answer space of a plural *wh*-question includes sum-based propositions while the answer space of a singular *wh*-question doesn’t. For example, in a context where two children came, the plural *wh*-question (77a) has an exhaustive true answer derived based on the sum of two children, $a \oplus b$. However, in the same context, the singular *wh*-question (77b) doesn’t have an exhaustive true answer, which violates the exhaustivity presupposition.

(77) (Among the children under consideration, only Andy and Billy came. The speaker knows that multiple children came, but she doesn’t know who they are.)

- a. Which children came? $\{\lambda w.\text{came}_w(a), \lambda w.\text{came}_w(b), \lambda w.\text{came}_w(a \oplus b)\}$
b. # Which child came? $\{\lambda w.\text{came}_w(a), \lambda w.\text{came}_w(b)\}$

Dayal (1996) is focused on the first-order interpretations of *wh*-questions. As for the higher-order interpretations, clearly singular *which*-phrases cannot quantify over Boolean conjunctions; otherwise (77b) would have an exhaustive true answer $\lambda w.\text{came}_w(a) \wedge \text{came}_w(b)$ formed out of $a^\uparrow \cap b^\uparrow$ and would satisfy the exhaustivity presupposition.³¹

³¹The same idea applies if we follow Fox (2018, 2020) and assume that a *wh*-domain may include higher-order pluralities. The uniqueness effects of singular *wh*-questions argue that singular *which*-phrases cannot quantify over higher-order pluralities; otherwise (77b) would have an exhaustive true answer based on the higher-order plurality $\{\{a\}, \{b\}\}$.

Higher-order *wh*-quantification exhibits a disjunction–conjunction asymmetry (Xiang 2021b): the higher-order quantification domain of a singular *which*-phrase includes Boolean disjunctions, even though it doesn't include Boolean conjunctions. For example in (78), the disjunctive answer to a singular *should*-question may be interpreted under the scope of the universal modal *should*. The narrow-scope interpretation of the disjunctive answer argues that the *wh*-domain of *which textbook* contains Boolean disjunctions over atomic books.

- (78) Which textbook should I use for this class?
Heim & Kratzer or Meaning & Grammar. (The choice is up to you.)

The disjunction–conjunction asymmetry is also manifested in singular *can*-questions. Unlike in (79), the MA answer to the singular *can*-question in (80) can be expressed as an elided FC-disjunction, but not as an elided conjunction.

- (79) What can I use for this class?
 a. *Heim & Kratzer or Meaning & Grammar.*
 b. *Heim & Kratzer and Meaning & Grammar.*
- (80) Which textbook can I use for this class?
 a. *Heim & Kratzer or Meaning & Grammar.*
 b. *# Heim & Kratzer and Meaning & Grammar.*

The following discussion on uniqueness will take disjunctive answers into account. These answers are especially important for local-uniqueness effects in modalized questions.

In addition to singular *which*-phrases, numeral-modified *which*-phrases also trigger a uniqueness presupposition, as exemplified in (81). The same analysis applies: *which two children* may quantify over pluralities of two children (e.g., $a \oplus b$) and the Boolean disjunctions of these pluralities (e.g., $(a \oplus b)^\uparrow \cup (c \oplus d)^\uparrow$), but not their sums or Boolean conjunctions or the higher-order pluralities of these pluralities (for more details, see Xiang 2021b). In (81b), for example, if *which two children* could quantify over Boolean conjunctions, the question would admit a conjunctive answer expressing the composition of two teams, derived from $(a \oplus b)^\uparrow \cap (c \oplus d)^\uparrow$, contrary to fact.

- (81) a. Which two children came? \rightsquigarrow *Only two of the children came.*
 b. Which two children formed a team? \rightsquigarrow *Only one pair of the children formed any team.*

5.1.2. Consequences of abandoning Dayal's exhaustivity presupposition

Dayal's exhaustivity presupposition requires a question to have an exhaustive true answer. In contrast, the max-informativity-based definition of answerhood adopted from Fox predicts that a MS interpretation is only available in cases where Dayal's exhaustivity presupposition is not trivially satisfied: a question has a MS interpretation if and only if this question can have multiple max-informative true answers (Sect. 4.1.2). This is where the dilemma arises: Dayal's exhaustivity presupposition explains uniqueness, but it is too strong to allow for MS; meanwhile Fox's concept of answerhood paves a way for MS, but is too weak to explain uniqueness.

Despite the status our model accords to MS, abandoning Dayal's exhaustivity presupposition would be too costly. In addition to explaining the uniqueness effects in questions, this exhaustivity presupposition has played an important role in accounting for a variety of linguistic phenomena, such as the negative island effects in degree constructions (Fox and Hackl 2007; Spector and Abrusán

2011; Abrusán 2014) and the maximality effects of definites (von Fintel et al. 2014). Moreover, it can rule out many unwanted question interpretations; without further restrictions, abandoning this presupposition and applying Fox’s concept of answerhood would let these unwanted interpretations be rebranded as acceptable MS interpretations. The following lays out three such problematic cases.

Case 1: Questions with a uniqueness presupposition. For a singular *wh*-question, the individual answers are mutually independent, and the answer space contains no plural or conjunctive answer stronger than these individual answers. Hence, without further restrictions, applying Fox’s concept of answerhood predicts a MS interpretation rather than a uniqueness effect. This problem also applies to numeral-modified *wh*-questions and alternative questions.

- (82) *Which child came? / Which two children formed a team? / Did you invite Andy, Billy, or Cindy?*
- a. With Dayal’s exhaustivity presupposition: MA with uniqueness ✓
 - b. Without Dayal’s exhaustivity presupposition: MS without uniqueness ✗

Case 2: *Wh*-questions with a non-distributive predicate. For *wh*-questions with a non-distributive predicate (e.g., *form a team, solve a problem together, be worth \$10 in total*), the individual answers are logically independent even if the *wh*-phrase is number-neutral or a bare plural. As seen in Sect. 4.2.3, for such a question, the answer space is closed under conjunction in the higher-order interpretation but not in the first-order interpretation. Hence, without further restrictions, applying Fox’s concept of answerhood predicts a MS interpretation for the first-order *wh*-quantification.

- (83) *Which children formed a team?*
- a. With Dayal’s exhaustivity presupposition: MA without uniqueness ✓
 - b. Without Dayal’s exhaustivity presupposition: Ambiguous between MS and MA ✗

Case 3: *Wh*-questions with an existential indefinite. *Wh*-questions with an existential indefinite have choice interpretations and individual interpretations. Thus the choice interpretation in (84a) requests the specification of a boy-movie(s) pair, whereas the individual interpretation in (84b) requests the specification of some movies, but not the names of the boys who watched these movies.

- (84) (Among the boys under consideration, Andy watched *Ironman* and *Spiderman*, and Billy watched only *Hulk*. Clark didn’t watch any movies.)
- Which movie or movies did one of the boys watch?
- a. ‘Name any/one boy *x*, and then tell me: Which movie(s) did *x* watch?’ (Choice)
 - i. Andy watched *Ironman* and *Spiderman*.
 - ii. Billy watched *Hulk*.
 - b. ‘Which movie(s) *y* is/are such that one of the boys watched *y*?’ (Individual)
 - i. One (of the boys) watched *Ironman* and *Spiderman*, and one watched *Hulk*.
 - ii. # One (of the boys) watched *Hulk*.

Existing literature is focused on the choice interpretation. A few works on MS (George 2011; Fox 2013; Nicolae 2013; a.o.) treat the choice interpretation as a MS interpretation and claim that existential indefinites can license MS. However, a number of empirical distinctions argue that choice interpretations and MS interpretations should be treated separately (Appendix A). Here, let’s switch

our focus to the individual interpretation. In this interpretation, the indefinite takes scope within the local IP. This raises a problem similar to what we've seen in questions with a collective predicate. In principle, in analogy to *can*-questions, questions with an existential indefinite could have an LF involving local exhaustification, as in (85a). In the context given above, the true answers yielded by this LF are those in (85b), both of which are max-informative. However, as seen in (84b-ii), such non-exhaustive individual answers are deviant. This problem also applies to the LF in (86a) for the higher-order *wh*-quantification.

- (85) a. $[_{CP} \text{ which-movie(s)} \lambda y_e [_{IP} \text{ one-boy } \lambda x_e [O_C [_{VP} x \text{ watched } y_{[+v]}]]]]$
 b. $\{ \lambda w. \exists x [\text{boy}_w(x) \wedge O_C[\text{watch}_w(x, i \oplus s)]] , \lambda w. \exists x [\text{boy}_w(x) \wedge O_C[\text{watch}_w(x, h)]] \}$
- (86) a. $[_{CP} \text{ which-movie(s)} \lambda \pi_{(et,t)} [_{IP} \text{ one-boy } \lambda x_e [\pi \lambda y_e [O_C [_{VP} x \text{ watched } y_{[+v]}]]]]]]$
 b. $\left\{ \begin{array}{l} \lambda w. \exists x [\text{boy}_w(x) \wedge O_C[\text{watch}_w(x, i \oplus s)]] , \lambda w. \exists x [\text{boy}_w(x) \wedge O_C[\text{watch}_w(x, h)]] \\ \lambda w. \exists x [\text{boy}_w(x) \wedge [O_C[\text{watch}_w(x, i \oplus s)] \vee O_C[\text{watch}_w(x, h)]] \end{array} \right\}$

The contrast between (84b-i) and (84b-ii) argues that the LFs in (85a) and (86a) are deviant; the LF for the individual interpretation should be like (87a), the answer space of which includes also answers formed out of Boolean coordinations and is closed under conjunction, as shown in (87b). However, once Dayal's exhaustivity presupposition is abandoned, it is unclear what constraints can rule out (85) and (86) and force (87).

- (87) a. $[_{CP} \text{ which-movie(s)} \lambda \pi_{(et,t)} [_{IP} \pi \lambda y_e [\text{one-boy } \lambda x_e [O_C [_{VP} x \text{ watched } y_{[+v]}]]]]]]$
 b. $\left\{ \begin{array}{l} \lambda w. \exists x [\text{boy}_w(x) \wedge O_C[\text{watch}_w(x, i \oplus s)]] \wedge \exists x [\text{boy}_w(x) \wedge O_C[\text{watch}_w(x, h)]] \\ \lambda w. \exists x [\text{boy}_w(x) \wedge O_C[\text{watch}_w(x, i \oplus s)]] , \lambda w. \exists x [\text{boy}_w(x) \wedge O_C[\text{watch}_w(x, h)]] \\ \lambda w. \exists x [\text{boy}_w(x) \wedge O_C[\text{watch}_w(x, i \oplus s)]] \vee \exists x [\text{boy}_w(x) \wedge O_C[\text{watch}_w(x, h)]] \end{array} \right\}$

5.1.3. A challenge to Dayal's exhaustivity presupposition: local uniqueness

As mentioned, Hirsch and Schwarz (2020) novelly observe that the uniqueness inference in a singular \diamond -question can take scope below the existential modal. For example, the following questions are acceptable in a multiple-choice context, where each choice involves a single letter, either 'a' or 'r'. As Hirsch and Schwarz argue, since Dayal's exhaustivity presupposition is applied to the answer space as a whole, it cannot account for these local-uniqueness inferences.

- (88) a. Which letter could we add to *fo_m* (to form a word)?
 A or r. (Intended: 'The unique letter that we add to *fo_m* could be a and could be r.')
- b. Which letter could be missing in *fo_m*?
 A or r. (Intended: 'The unique letter missing in *fo_m* could be a and could be r.')

According to Hirsch and Schwarz, local uniqueness is available regardless of modal flavor: in the \diamond -questions in (88), the modal verb *could* is a priority modal in (88a) and an epistemic modal in (88b). Below, I further argue that local uniqueness is available regardless of the modal force. For example, the singular *have to*-question in (89) implies that we cannot assign more than one chapter to the students.

- (89) Q: Which chapter do we have to assign to the students?
 \rightsquigarrow We are not allowed to assign more than one chapter to the students.
- A: Chapter 1 or chapter 2, either is good. $(\Box(\phi_{c_1} \vee \phi_{c_2}))$

Dayal’s exhaustivity presupposition cannot account for the local uniqueness inference in (89). To see why, recall that the singular *which*-phrase may only range over atomic entities and their Boolean disjunctions. With two chapters c_1 and c_2 in the discourse domain, the answer space yielded in a narrow-scope higher-order interpretation is the one given in (90a). Next, assume the scenario in (90b), where local uniqueness is violated: among the three w -accessible worlds w_1, w_2, w_3 , we assign only c_1 in w_1 , only c_2 in w_2 , but both c_1 and c_2 in w_3 . In this scenario, the answer space has only one true member, $\Box(\phi_{c_1} \vee \phi_{c_2})$, and hence Dayal’s exhaustivity presupposition is not violated.

- (90) Which chapter do we have to assign to the students?
 $[\text{CP which-chapter } \lambda\pi_{\langle et, t \rangle} [\text{IP have to } [\pi \lambda x_e [\text{VP we assign } x \text{ to the students }]]]]$
 a. $Q = \{\Box\phi_{c_1}, \Box\phi_{c_2}, \Box(\phi_{c_1} \vee \phi_{c_2})\}$ (ϕ_x abbreviates ‘we assign x to the students’)
 b. $M_w = \{w_1, w_2, w_3\}$, $\text{assign} = [w_1 \rightarrow \{c_1\}, w_2 \rightarrow \{c_2\}, w_3 \rightarrow \{c_1, c_2\}, \dots]$

5.2. Two alternative accounts

5.2.1. Partition by exhaustification (Fox 2018, 2020)

Fox (2018, 2020) presents a solution to the dilemma between uniqueness and MS. These works cover a number of issues, including also the negative island effects of higher-order *wh*-quantification. The following will focus on those aspects that are directly relevant to the dilemma between MS and uniqueness.

Fox first re-writes Dayal’s definition of answerhood in terms of exhaustification; his version is given in (91): for a world w and a Hamblin set Q , applying $\text{ANS}_{\text{Dayal}}$ returns the unique proposition p in Q such that the exhaustification of p relative to Q is true in w ; this application is defined only if such a proposition exists. In this definition, the O -operator is the traditional exhaustification operator: it affirms the prejacent and negates all the alternatives that are not entailed by this prejacent.

$$(91) \text{ANS}_{\text{Dayal}}(w)(Q) = \exists p \in Q [O_Q(p)(w) = 1]. p \in Q [O_Q(p)(w) = 1] \\ \text{where } O_Q(p) := \lambda w'. p(w') = 1 \wedge \forall q \in Q [p \not\subseteq q \rightarrow q(w') = 0]$$

Next, Fox modifies the definition by replacing the traditional O -operator with one that may trigger FC. He assumes the exhaustivity operator $O^{\text{IE+II}}$ from Bar-Lev and Fox 2020. As defined below, this operator negates the innocently excludable (IExcl)-alternatives (called ‘*Innocent Exclusion* (IE)’) and affirms the innocently includable (IIIncl)-alternatives (called ‘*Innocent Inclusion* (II)’).

$$(92) O_C^{\text{IE+II}} = \lambda p \lambda w. \underbrace{\forall q \in C [q \in \text{IEXCL}(p, C) \rightarrow q(w) = 0]}_{\text{Innocent Exclusion (IE)}} \wedge \underbrace{\forall r \in C [r \in \text{IIINCL}(p, C) \rightarrow r(w) = 1]}_{\text{Innocent Inclusion (II)}} \\ \text{a. IEXCL}(p, C) \\ = \bigcap \{A \subseteq C \mid A \text{ is a maximal subset of } C \text{ s.t. } \{\neg q \mid q \in A\} \cup \{p\} \text{ is consistent}\} \\ \text{b. IIINCL}(p, C) \\ = \bigcap \{B \subseteq C \mid B \text{ is a maximal subset of } C \text{ s.t. } B \cup \{p\} \cup \{\neg q \mid q \in \text{IEXCL}(p, C)\} \text{ is consistent}\}$$

Let’s see how the $O^{\text{IE+II}}$ -operator derives FC and predicts its distribution. In (93), applying $O^{\text{IE+II}}$ to the modalized disjunction $\Diamond(\phi_a \vee \phi_b)$ yields a scalar implicature $\neg\Diamond(\phi_a \wedge \phi_b)$ by innocent exclusion, as well as a FC inference $\Diamond\phi_a \wedge \Diamond\phi_b$ by innocent inclusion. In contrast, in (94), applying $O^{\text{IE+II}}$ to a plain disjunction doesn’t yield FC, because the disjuncts are not I-includable: affirming both disjuncts yields an inference conflicting with the negation of the I-excludable alternative (viz., $\phi_a \wedge \phi_b$ and $\neg(\phi_a \vee \phi_b)$ are contradictory). Hence, $O^{\text{IE+II}}$ may yield a FC inference for a disjunctive

sentence only if the alternative set of this sentence is not closed under conjunction, such as in cases where the disjunction takes scope below an existential modal.

- (93) Let $p = \diamond(\phi_a \vee \phi_b)$ and $C = \{\diamond\phi_a, \diamond\phi_b, \diamond(\phi_a \vee \phi_b), \diamond(\phi_a \wedge \phi_b)\}$; then we have:
- $\text{IE}_{\text{EXCL}}(p, C) = \{\diamond(\phi_a \wedge \phi_b)\}$ and $\text{II}_{\text{NCL}}(p, C) = \{\diamond\phi_a, \diamond\phi_b, \diamond(\phi_a \vee \phi_b)\}$
 - $O_C^{\text{IE}+\text{II}}(p) = \underbrace{\neg\diamond(\phi_a \wedge \phi_b)}_{\text{IE}} \wedge \underbrace{[\diamond\phi_a \wedge \diamond\phi_b \wedge \diamond(\phi_a \vee \phi_b)]}_{\text{II}} = \neg\diamond(\phi_a \wedge \phi_b) \wedge \diamond\phi_a \wedge \diamond\phi_b$
- (94) Let $p = \phi_a \vee \phi_b$ and $C = \{\phi_a, \phi_b, \phi_a \vee \phi_b, \phi_a \wedge \phi_b\}$; then we have:
- $\text{IE}_{\text{EXCL}}(p, C) = \{\phi_a \wedge \phi_b\}$ and $\text{II}_{\text{NCL}}(p, C) = \{\phi_a \vee \phi_b\}$
 - $O_C^{\text{IE}+\text{II}}(p) = \underbrace{\neg[\phi_a \wedge \phi_b]}_{\text{IE}} \wedge \underbrace{[\phi_a \vee \phi_b]}_{\text{II}}$

Dayal's exhaustivity presupposition can now be modified to the following:

- (95) *Modified exhaustivity presupposition* (after Fox 2020)
 For any question with a Hamblin set Q , the application of the answerhood operator to Q is defined in w only if there is a proposition p in Q such that $O_Q^{\text{IE}+\text{II}}(p)$ is true in w .

Just like Dayal's exhaustivity presupposition, this modified presupposition is trivially satisfied if the answer space is closed under conjunction. However, this modified presupposition can also be trivially satisfied in cases where the answer space contains disjunctions that can be strengthened into FC statements via the application of $O^{\text{IE}+\text{II}}$. This property allows for MS.

To see explicitly how this analysis allows for MS, consider the *can*-question in (96). With three relevant individuals a, b, c , the answer space of the question is as laid out in (96a).³² Applying $O_Q^{\text{IE}+\text{II}}$ point-wise to this answer space returns a set of mutually exclusive propositions as in (96b), referred to as 'the partition induced by Q '. The modified exhaustivity presupposition is satisfied in any world that is in the union of this partition, including worlds with multiple possible chair persons.

- (96) Who can chair the committee alone?
- $Q = \left\{ \begin{array}{l} \diamond\phi_a, \diamond\phi_b, \diamond\phi_c \\ \diamond(\phi_a \vee \phi_b), \diamond(\phi_b \vee \phi_c), \diamond(\phi_a \vee \phi_c) \\ \diamond(\phi_a \vee \phi_b \vee \phi_c) \end{array} \right\}$
 $(\phi_x \text{ abbreviates 'x chairs the committee alone'})$
 - $\{O_Q^{\text{IE}+\text{II}}(p) \mid p \in Q\}$
 $= \left\{ \begin{array}{l} \diamond\phi_a \wedge \neg\diamond\phi_b \wedge \neg\diamond\phi_c, \diamond\phi_b \wedge \neg\diamond\phi_a \wedge \neg\diamond\phi_c, \diamond\phi_c \wedge \neg\diamond\phi_a \wedge \neg\diamond\phi_b \\ \diamond\phi_a \wedge \diamond\phi_b \wedge \neg\diamond\phi_c, \diamond\phi_b \wedge \diamond\phi_c \wedge \neg\diamond\phi_a, \diamond\phi_a \wedge \diamond\phi_c \wedge \neg\diamond\phi_b \\ \diamond\phi_a \wedge \diamond\phi_b \wedge \diamond\phi_c \end{array} \right\}$

This partition-by-exhaustification analysis also avoids over-generating MS interpretations for questions with a non-distributive predicate. For example, in a multiple-team scenario, question (97) (see also (52)) satisfies the exhaustivity presupposition only if its answer space is closed under conjunction. Such an answer space can be created if the *wh*-phrase ranges over higher-order pluralities

³²For Fox's (2018, 2020) analysis, it is important to exclude the narrow-scope conjunctive answers (e.g., $\diamond(\phi_a \wedge \phi_b)$) from the answer space of a MS question. Fox assumes that such answers are 'pruned'. For simplicity, I here demonstrate Fox's analysis using an example with the predicate *chair the committee alone*, for which we can ignore such conjunctive answers.

(as assumed in Fox 2018, 2020, discussed in Sect. 4.2.3) or Boolean conjunctions (as assumed in my proposal).³³

(97) Q: ‘Which children formed a team?’

A: ‘Andy and Billy formed a team, and Cindy and Danny formed a team.’

What about uniqueness? At first sight, the partition-by-exhaustification analysis should preserve the merits of Dayal’s exhaustivity presupposition in explaining uniqueness: if a question doesn’t have FC answers, the modification to Dayal’s exhaustivity presupposition shouldn’t affect the predictions. However, as I argue below, this problem remains unsolved because the $O^{\text{IE}+\text{II}}$ -operator over-generates FC readings for disjunctive answers to singular *wh*-questions.

For an illustration of the uniqueness problem, consider the singular *wh*-question in (98). With two boys a and b in the discourse domain, the answer space is Q_{FO} in the first-order interpretation and Q_{HO} in the higher-order interpretation.³⁴ In (98b), applying point-wise exhaustification to Q_{FO} induces a partition that covers only the worlds where exactly one of the boys came, which does nicely predict uniqueness. However, in the higher-order interpretation, exhaustifying the non-modalized disjunctive answer $\phi_a \vee \phi_b$ with the $O^{\text{IE}+\text{II}}$ -operator would strengthen it into a FC statement: distinct from the alternative set C in (94), here Q_{HO} doesn’t contain the conjunction $\phi_a \wedge \phi_b$, and thus the individual answers ϕ_a and ϕ_b are I-includable w.r.t. the disjunctive answer $\phi_a \vee \phi_b$. Hence in (98c), the partition induced by Q_{HO} covers also the worlds where both boys came, which therefore predicts no uniqueness effect. In short, the uniqueness effect remains unexplained because the $O^{\text{IE}+\text{II}}$ -operator over-generates FC answers for singular *wh*-questions with a higher-order interpretation.

(98) Which boy came?

- a. $Q_{\text{FO}} = \{\phi_a, \phi_b\}$ and $Q_{\text{HO}} = \{\phi_a, \phi_b, \phi_a \vee \phi_b\}$ (ϕ_x abbreviates ‘ x came’)
 - b. $\{O_{Q_{\text{FO}}}^{\text{IE}+\text{II}}(p) \mid p \in Q_{\text{FO}}\} = \{\phi_a \wedge \neg\phi_b, \neg\phi_a \wedge \phi_b\}$ (Prediction: uniqueness ✓)
 - c. $\{O_{Q_{\text{HO}}}^{\text{IE}+\text{II}}(p) \mid p \in Q_{\text{HO}}\} = \{\phi_a \wedge \neg\phi_b, \neg\phi_a \wedge \phi_b, \phi_a \wedge \phi_b\}$ (Prediction: no uniqueness ✗)
- [N.B.: $\text{IEXCL}(\phi_a \vee \phi_b, Q_{\text{HO}}) = \emptyset$ and $\text{INCL}(\phi_a \vee \phi_b, Q_{\text{HO}}) = \{\phi_a \vee \phi_b, \phi_a, \phi_b\}$;
therefore: $O_{Q_{\text{HO}}}^{\text{IE}+\text{II}}(\phi_a \vee \phi_b) = \phi_a \wedge \phi_b$.]

The $O^{\text{IE}+\text{II}}$ -operator also over-predicts FC for disjunctions embedded under an existential indefinite, which in turn over-predicts a MS interpretation for questions with an existential indefinite. As argued in Sect. 5.1.2, the answer space of (99) couldn’t be the one given in (99a), since this would allow for a non-exhaustive individual interpretation.³⁵ Applying $O_Q^{\text{IE}+\text{II}}$ point-wise to this answer space yields the partition in (99b), which includes worlds where two movies were watched, each by a different boy.

³³The analysis proposed in Fox 2018, 2020 is more complex than what is presented here. To account for the negative island effects in higher-order interpretations, Fox also argues for a *non-vacuity* principle:

(i) *Non-vacuity* (Fox 2018, 2020)

Every proposition p in Q is such that the exhaustification of p is identical to a cell in the partition induced by Q .

This principle predicts that simple *wh*-questions like *Who left?* and *Which children formed a team?* do not have higher-order interpretations: in a higher-order interpretation, the answer space of a simple *wh*-question contains plain disjunctions like $\phi_a \vee \phi_b$, which cannot be paired with a partition cell by exhaustification. To account for the data in (97), Fox further assumes that the quantification domain of *which children* includes higher-order pluralities such as $\{\{a, b\}, \{c, d\}\}$. With this assumption, the answer space yielded in a first-order interpretation is closed under conjunction.

³⁴As mentioned, singular *which*-phrases may range over Boolean disjunctions (e.g., $a^\uparrow \cap b^\uparrow$), but not over Boolean conjunctions (e.g., $a^\uparrow \cup b^\uparrow$) or higher-order pluralities (e.g., $\{\{a\}, \{b\}\}$).

³⁵To be consistent with Fox’s analysis, I here remove the local exhaustifier assumed in my own proposal and prune the narrow-scope conjunctive answer $\exists x[\phi_{\langle x, m_1 \rangle} \wedge \phi_{\langle x, m_2 \rangle}]$. These changes do not affect the illustration of this issue.

(99) Which movie or movies did one of the boys watch? (Individual)

$$a. Q = \left\{ \exists x.\phi_{\langle x, m_1 \rangle}, \exists x.\phi_{\langle x, m_2 \rangle}, \exists x[\phi_{\langle x, m_1 \rangle} \vee \phi_{\langle x, m_2 \rangle}] \right\}$$

($\phi_{\langle x, m \rangle}$ abbreviates ‘boy x watched m ’)

$$b. \{O_Q^{\text{IE}+\text{II}}(p) \mid p \in Q\}$$

$$= \left\{ \exists x.\phi_{\langle x, m_1 \rangle} \wedge \neg \exists x.\phi_{\langle x, m_2 \rangle}, \exists x.\phi_{\langle x, m_2 \rangle} \wedge \neg \exists x.\phi_{\langle x, m_1 \rangle}, \exists x.\phi_{\langle x, m_1 \rangle} \wedge \exists x.\phi_{\langle x, m_2 \rangle} \right\}$$

[N.B.: $\text{IEXCL}(\exists x[\phi_{\langle x, m_1 \rangle} \vee \phi_{\langle x, m_2 \rangle}], Q) = \emptyset$ and $\text{II}_{\text{NCL}}(\exists x[\phi_{\langle x, m_1 \rangle} \vee \phi_{\langle x, m_2 \rangle}], Q) = Q$;
therefore: $O_Q^{\text{IE}+\text{II}}(\exists x[\phi_{\langle x, m_1 \rangle} \vee \phi_{\langle x, m_2 \rangle}]) = \exists x.\phi_{\langle x, m_1 \rangle} \wedge \exists x.\phi_{\langle x, m_2 \rangle}$]

In short, since the $O^{\text{IE}+\text{II}}$ -operator predicts no difference between existential modals and existential indefinites in licensing FC, the partition-by-exhaustification analysis predicts no difference between existential modals and existential quantifiers in licensing MS.

To sum up, the partition-by-exhaustification analysis of Fox (2018, 2020) allows for MS; however, due to the particular choice of the $O^{\text{IE}+\text{II}}$ -operator, this account remains insufficient in predicting uniqueness effects and in avoiding over-generation of MS. One way to fix these problems would be to use a different method to derive partition and FC. For example, the partition can be induced by applying both the IE-based exhaustification operator O^{IE} and the anti-exhaustification operator DOU to each propositional answer. The $O^{\text{IE}+\text{II}}$ -operator and the DOU -operator predict different distributions of FC. As discussed above, exhaustifying a disjunction-containing sentence by $O^{\text{IE}+\text{II}}$ yields FC if and only if the alternative set of this sentence isn’t closed under conjunction. The $O^{\text{IE}+\text{II}}$ -based analysis predicts FC for $\diamond(\phi \vee \psi)$ as well as for $\exists x[\phi_{\langle x, a \rangle} \vee \phi_{\langle x, b \rangle}]$, but not for wide-scope disjunctions including $\diamond\phi \vee \diamond\psi$. It also predicts FC for the plain disjunction $\phi \vee \psi$ if the conjunctive alternative $\phi \wedge \psi$ is pruned. In contrast, the DOU -based analysis, which accounts for the distribution of universal FC based on the RelExcl presupposition, predicts universal FC only for \diamond -disjunctions like $\diamond(\phi \vee \psi)$ and $\diamond\phi \vee \diamond\psi$.

Fox’s (2018, 2020) account doesn’t deal with local-uniqueness effects. For a recent analysis of local uniqueness based on partition and local exhaustification, see Kobayashi and Rouillard 2021.

5.2.2. Presuppositional *which* (Hirsch and Schwarz 2020)

As mentioned before, to account for local-uniqueness effects in \diamond -questions, Hirsch and Schwarz (2020) propose, in line with Rullmann and Beck (1998), that the uniqueness presupposition of a singular *wh*-question stems from the lexical meaning of the determiner *which*. They define *which* as in (100) and assume that *which* is interpreted within the question nucleus.³⁶

$$(100) \llbracket \textit{which} \rrbracket = \lambda x_e \lambda m_{\langle e, st \rangle} \lambda g_{\langle e, st \rangle} \lambda w : \exists ! y [f(y)(w) \wedge g(y)(w)]. \underline{f(x)(w)} \wedge g(x)(w)$$

Hirsch and Schwarz further argue that global/local uniqueness arises if the *which*-phrase takes scope above/below the existential modal, as exemplified in (101). Here each propositional answer carries a global/local uniqueness presupposition (underlined). For the local-uniqueness interpretation (101b), if none of the accessible worlds satisfies uniqueness, the question has no true answer and is deviant.

³⁶The main texts of Uegaki 2018, 2021 also define *which* with a uniqueness presupposition. The primary goal of Uegaki’s analysis is to account for the projection behavior of the uniqueness presupposition in question embeddings. However, for his core idea it only matters that the uniqueness presupposition is carried by each answer individually (as opposed to by the answer set as a whole); it does not matter to him how this presupposition is compositionally derived. In Uegaki 2021: Appendix A, he sketches out an alternative analysis which assumes that the uniqueness presupposition is assigned to each of the propositional answers after the application of the ANS -operator. My proposal in Sect. 6 can quite easily be made compatible with this analysis; for details, see fn. 41.

- (101) Which letter could we add to fo_m ?
- a. *which* \gg *could*: global uniqueness
 - i. $[\text{CP ? } \lambda 1 [\text{IP } [\text{DP } \text{which } t_1 \text{ letter}] \lambda 2 [\text{ could } [\text{VP } \text{we add } t_2 \text{ to } fo_m]]]]$
 - ii. $\{\lambda w : \exists! y [\diamond_w [\lambda w'. \text{letter}_{w'}(y) \wedge \text{add}_{w'}(y)]] . \diamond_w [\lambda w'. \text{letter}_{w'}(a) \wedge \text{add}_{w'}(a)] \mid x \in D_e\}$
 - b. *could* \gg *which*: local uniqueness
 - i. $[\text{CP ? } \lambda 1 [\text{IP } \text{could } [[\text{DP } \text{which } t_1 \text{ letter}] \lambda 2 [\text{VP } \text{we add } t_2 \text{ to } fo_m]]]]$
 - ii. $\{\lambda w : \diamond_w [\lambda w'. \exists! y [\text{letter}_{w'}(y) \wedge \text{add}_{w'}(y)]] . \diamond_w [\lambda w'. \text{letter}_{w'}(a) \wedge \text{add}_{w'}(a)] \mid x \in D_e\}$

The above example concerns singular *wh*-questions. To allow their account to apply to plural *wh*-questions, Hirsch and Schwarz re-define the semantics of *which* as follows:

$$(102) \llbracket \text{which} \rrbracket = \lambda x_e \lambda m_{\langle e, st \rangle} \lambda g_{\langle e, st \rangle} \lambda w : \exists y [f(y)(w) \wedge g(y)(w) \wedge \forall z [f(y)(w) \wedge g(y)(w) \rightarrow z \leq y]]. \\ f(x)(w) \wedge g(x)(w)$$

In this definition, the presupposition of *which* is trivially satisfied if and only if the intersection between the *wh*-domain and the extension of the predicate that the *which*-phrase combines with is closed under sum. For example, this presupposition is trivially satisfied in *Which students came?*: for any world w , the set of students who came in w (viz., $\llbracket \text{students} \rrbracket^w \cap \llbracket \text{came} \rrbracket^w$) is closed under sum. In contrast, *Which student came?* is subject to uniqueness: the set of atomic students who came in w (viz., $\llbracket \text{student} \rrbracket^w \cap \llbracket \text{came} \rrbracket^w$) is closed under sum if and only if exactly one student came in w .

Compared with the analysis of uniqueness based on Dayal's exhaustivity presupposition, the presuppositional-*which* account has advantages in deriving local-uniqueness interpretations. Moreover, this account is compatible with Fox's answerhood and permits MS.

However, despite these advantages, the presuppositional-*which* account faces many problems. First, this account only deals with uniqueness effects; it cannot avoid over-generating MS interpretations for *wh*-questions with a collective predicate or with an existential indefinite.

Second, this account predicts a local-uniqueness effect for a singular *can*-question, but not for the MS answers to this question: it doesn't require the answer chosen by the addressee to be one that satisfies local uniqueness. To see this problem more concretely, consider (103). Given the existence of the option to assign a single chapter, Hirsch and Schwarz's account predicts that the local-uniqueness requirement of the embedded singular *can*-question *which chapter we can assign to the students* is satisfied, and further, that any true answer that specifies a single chapter is a good MS answer to this question. However, the continuation in (103) is clearly unacceptable: the uniqueness requirement is not satisfied in worlds where we assign "chapter 3".³⁷

- (103) (The book under consideration has three chapters. The speaker, who is a TA of the class, has been informed by the instructor that they could assign either chapter 1, or chapter 2, or both chapters 2 and 3 to the students next week.)
- # I know which chapter we can assign to the students (next week) ..., chapter 3.

Third, the local-uniqueness inference predicted by this account is existential, which is too weak. For example, for the embedded question *which chapter we can assign to the students*, the predicted local-uniqueness inference is read as "We are allowed to assign a single chapter to the students", not

³⁷Eagle-eyed readers might find the above statement odd even before the answer continuation "chapter 3" is given. This oddness is due to the problem we turn to next, which applies to the embedded question itself and is independent of the answer continuation.

“We are *only* allowed to assign a single chapter to the students”. However, (104a) is marginal, despite the fact that uniqueness is satisfied in all of the accessible worlds where chapter 1 is assigned and in some of the accessible worlds where chapter 2 is assigned. In contrast, it is more natural to express the intended meaning with the stressed modifier *SIN*gle, as in (104b), or by using the number-neutral word *what*, as in (104c). (For an explanation of the contrast between (104a) and (104b), see Sect. 6.3.3.)

- (104) (As in (103).)
- a. ? I know which chapter we can assign to the students next week ...,
chapter 1/ chapter 2/ chapter 1 or chapter 2.
 - b. I know which SINgle chapter we can assign to the students next week ...,
chapter 1/ chapter 2/ chapter 1 or chapter 2.
 - c. I know what we can assign to the students next week.

Contrary to Hirsch and Schwarz 2020, I argue that the uniqueness requirement in a modalized singular *wh*-question needs to be satisfied in *every* accessible world which alone verifies a true answer. Dialogue (105) illustrates this requirement. In an informal survey I conducted, 9 out of 11 native speakers reported that TA1’s utterance sounded unnatural or sly — it seemed to them that TA1 was unintentionally assuming or intentionally suggesting that they should assign at most one paper next week, thereby modifying the instructor’s request. In comparison, dialogue (106) was deemed perfectly natural. According to my informants, TA2’s question clearly doesn’t imply that the reading for next week must be a journal article; what she were asking is simply a sub-question of the more general question ‘Which journal article or book chapter should we assign to the students next week?’, namely, ‘If we choose to assign a journal article, what could it be?’

- (105) Instructor: “We should assign one or two papers to the students each week.”
TA1: “Got it. Which paper could we assign to the students next week?” [Unnatural]
- (106) Instructor: “We should assign a journal article or a book chapter to the students each week.”
TA2: “Got it. Which journal article could we assign to the students next week?” [Natural]

Fourth, the presuppositional-*which* account under-predicts uniqueness effects for numeral-modified *wh*-questions that have a non-distributive predicate. Like singular *which*-phrases, numeral-modified *which*-phrases also trigger uniqueness (Xiang 2021b; see also Sect. 5.1.1). For example, the numeral-modified *wh*-question (107) expresses the same meaning as the singular *wh*-question in (107a) (‘singular’ in the sense that the noun *group* is singular), and it implies the uniqueness inference in (107b).

- (107) Which *N* students solved a problem together?
- a. ‘Which (unique) *N*-membered group of students solved a problem together?’
 - b. \rightsquigarrow *Only one N-membered group of students solved a problem together.*

In the scenario described in (108) below, the uniqueness inference of (108a) is false, while that of (108b) is true. This distinction causes the contrast in felicity between (108a) and (108b). Moreover, (108c) is also infelicitous: its uniqueness inference, namely that only one 2–3-membered set of students solved a problem together, is false.³⁸ In comparison, as seen in (108d), the uniqueness effect is absent when

³⁸According to my informants, (108c) is quite unnatural, aside from the deviance caused by the uniqueness effect. Some of the informants judged the use of the complex *wh*-expression ‘*which two or three NP*’ unnatural. The others accepted the use of this *wh*-expression (and judged (108d) grammatical and felicitous), but they found sentence (108c) ambiguous. Other than the

the predicate *solved a problem together* is replaced with its plural and semi-distributive counterpart *solved problems together*. The presuppositional semantics of *which* in (102) predicts the facts about uniqueness in (108a,b) but not the violation of uniqueness in (108c): in the described scenario, the intersection between the *wh*-domain and the extension of the predicate in (108c) is closed under sum.

- (108) (The students solved three problems in total: $a + b$ solved one, $b + c$ solved one, and $a + b + c$ solved one. The speaker knows how many students participated in solving each problem, but she doesn't know who these students are.)
- a. # Which two students solved a problem together?
 \rightsquigarrow Only one 2-membered group of students solved a problem together.
 $\llbracket \text{two students} \rrbracket^w \cap \llbracket \text{s.a.p.t.} \rrbracket^w = \{a \oplus b, b \oplus c\}$
 (Violation of uniqueness is predicted: the presupposition of *which* is false in w .)
 - b. Which three students solved a problem together?
 \rightsquigarrow Only one 3-membered group of students solved a problem together.
 $\llbracket \text{three students} \rrbracket^w \cap \llbracket \text{s.a.p.t.} \rrbracket^w = \{a \oplus b \oplus c\}$
 (Satisfaction of uniqueness is predicted: the presupposition of *which* is true in w .)
 - c. # Which two or three students solved a problem together?
 \rightsquigarrow Only one 2–3-membered group of students solved a problem together.
 $\llbracket \text{two or three students} \rrbracket^w \cap \llbracket \text{s.a.p.t.} \rrbracket^w = \{a \oplus b, b \oplus c, a \oplus b \oplus c\}$
 (Violation of uniqueness isn't predicted: the presupposition of *which* is true in w .)
 - d. Which two or three students solved problems together?
 (No uniqueness effect.)

More generally, Dayal's exhaustivity presupposition is concerned with the entailment relation of propositional answers, while the presupposition of *which* assumed in (102) is concerned with the 'part-of' relation of short answers. In *wh*-questions with a non-distributive predicate, the satisfaction of the part-of relation does not ensure the satisfaction of the entailment relation. The uniqueness implication observed in cases like (108c) argue that such uniqueness effects in questions are due to a constraint on propositions, not a constraint on individuals.

Last, in questions with multiple singular *which*-phrases, the uniqueness presupposition assumed for the higher /subject *which*-phrase is too strong to allow for a pair-list interpretation. In the pair-list interpretation (109b), there is a point-wise uniqueness requirement w.r.t. the movies under consideration, but no uniqueness requirement w.r.t. the boys. To allow for the observed non-uniqueness w.r.t. the subject domain, we are forced to assume a non-presuppositional semantics for the subject *which boy*, which clearly conflicts with Hirsch and Schwarz's explanation of uniqueness.

- (109) Which boy watched which movie?
- a. Single-pair interpretation: 'Which unique boy- x -movie- y pair is such that x watched y ?'
 'Andy watched *Spiderman*.'
 - b. Pair-list interpretation: '[Each boy watched at most one movie; tell me:] which boy- x -movie- y pairs are such that x watched y ?'
 'Andy watched *Ironman*, Billy watched *Spiderman*, Clark watched *Hulk*.'

intended meaning in (i-a), (108c) can be understood as in (i-b), where the disjunctive *or* takes wide scope and disjoins two *wh*-questions. This ambiguity makes the utterance (108c) quite inefficient in the conversation.

- (i) a. 'Which (unique) 2–3-membered group of students solved a problem together?'
- b. 'Which two students [...] or which three students solved a problem together?'

By contrast, point-wise uniqueness in pair-list interpretations can be derived based on Dayal’s exhaustivity presupposition. See Dayal 1996, 2017 for details, and see Fox 2012 and Xiang 2019, 2021a,c for two alternative accounts that assume Dayal’s exhaustivity presupposition.

6. Solving the dilemma: Relativized Exhaustivity

Section 5 has presented a dilemma: Dayal’s exhaustivity presupposition is incompatible with MS, but abandoning this condition would leave uniqueness effects unexplained and would over-generate MS interpretations for a variety of types of questions. Further, the presented observations on local uniqueness have argued that Dayal’s exhaustivity presupposition has a general problem in tackling modalized questions. This problem is independent of our considerations on MS: local uniqueness is observed in various modalized questions, regardless of modal flavor and modal force. Hence, we need an alternative to Dayal’s exhaustivity presupposition that correctly predicts the distribution of MS and uniqueness.

6.1. Relativized Exhaustivity

In a modalized singular *wh*-question, the uniqueness requirement appears to be ‘local’ if uniqueness is evaluated relative to the accessible worlds, as opposed to the anchor/utterance world. For example, for the modalized questions in (110a,b), local uniqueness says that in every accessible world at most one chapter is assigned to the students. This inference is more accurately stated as in (110c), where w is the utterance world, and M is the contextually determined modal base of *can/have to* which maps the utterance world to a set of accessible worlds.

- (110) a. Which chapter can we assign to the students?
 b. Which chapter do we have to assign to the students?
 c. ‘For every world w' in M_w such that we assign any chapters to the students in w' , we assign only one chapter to the students in w' .’

Further, in light of Dayal’s insight that uniqueness comes from exhaustivity, I argue that question interpretations do require exhaustivity, but in modalized questions the required exhaustivity is evaluated relative to the accessible worlds, not the utterance world.³⁹

How can we capture this idea in model-theoretic compositional semantics? Clearly, this cannot be achieved by changing the interpretation world from the utterance world to the worlds accessible to the utterance world: for the modalized *wh*-question *Who can/has to P?*, what we want is to evaluate the exhaustivity of the corresponding non-modalized question *Who P?* relative to the accessible worlds, not to evaluate the exhaustivity of the original question *Who can/has to P?* relative to the accessible worlds. Alternatively, as demonstrated below, the idea about the perspectival shift from the utterance world to the accessible worlds can be captured by substituting the modal base with singleton sets consisting of only one of the accessible worlds.

First, I redefine Dayal’s exhaustivity presupposition as in (111), where the semantic denotation of the question is sensitive to the modal base. Here $\llbracket Q \rrbracket_w^M$ ($:= \{\alpha \mid \alpha \in \text{Dom}(\llbracket Q \rrbracket^M) \wedge \llbracket Q \rrbracket^M(\alpha)(w) = 1\}$) abbreviates the set of short answers to Q that are true in w given the modal base M .

³⁹One might speculate that the relativization occurs as language users frequently shift the origo for the semantic interpretation. The most well-known phenomenon of perspectival shift is *deictic projection*, which occurs when the speakers ostensibly position themselves in a place and/or time other than the place and time of the utterance. Similarly, when evaluating properties like uniqueness and exhaustivity for a modalized question, the language users might shift their perspective for meaning evaluation from the utterance world to the accessible worlds where the state under discussion actually emerges.

(111) *Dayal's exhaustivity presupposition* (adapted from Dayal 1996)

Given a modal base M , a question Q is defined in w only if

$$\exists \alpha [\alpha \in \llbracket Q \rrbracket_w^M \wedge \forall \beta [\beta \in \llbracket Q \rrbracket_w^M \rightarrow \llbracket Q \rrbracket^M(\alpha) \subseteq \llbracket Q \rrbracket^M(\beta)]] \quad [\text{abbreviated as } \mathbf{DEP}(w, M, \llbracket Q \rrbracket)]$$

Next, I propose that questions are subject to a *Relativized Exhaustivity* (RelExh) condition, defined as in (112). This condition requires Dayal's exhaustivity presupposition to be satisfied relative to every modal base that introduces a singleton set of accessible worlds which verifies a true answer. (See Sect. 4.4.3 on Relativized Exclusivity (RelExcl) for definitions of relevant concepts.)⁴⁰

(112) *Relativized Exhaustivity*

Given a modal base M , a question Q is defined in w only if

$$\forall M'_{(s,st)} [|M'_w| = 1 \wedge M'_w \subseteq M_w \wedge \exists \alpha [\alpha \in \llbracket Q \rrbracket_w^M \wedge \alpha \in \llbracket Q \rrbracket_w^{M'}] \rightarrow \mathbf{DEP}(w, M', \llbracket Q \rrbracket)]$$

'For every modal base M' such that M'_w is a singleton subset of M_w , if M'_w verifies one of the true short answers to Q in w (viz., some true short answer to Q in w given M is also a true short answer to Q in w given M'), the interpretation of Q relative to M' satisfies Dayal's exhaustivity presupposition in w .' [abbreviated as $\mathbf{REP}(w, M, \llbracket Q \rrbracket)$]

The next two subsections will present the applications and predictions of this RelExh condition. This condition allows *can*-questions to have a MS interpretation (Sect. 6.2.1) and avoids over-generating MS interpretations for non-*can*-questions (Sect. 6.2.2). Moreover, it naturally accounts for the local-uniqueness effects in modalized questions (Sect. 6.3).

6.2. Predictions of RelExh on the distribution of MS

6.2.1. Permitting MS

To see how RelExh permits MS, consider the following MS interpretations of a *can*-question:

(113) Who can chair the committee? (ϕ_x abbreviates 'x chairs the committee')

a. $\llbracket Q_{\text{FO}} \rrbracket^M = \lambda x_e : \text{hmn}_@ (x). \exists w' \in M_w [w' \in O_C \phi_x]$ (First-order MS)

b. $\llbracket Q_{\text{HO}} \rrbracket^M = \lambda \pi_{ett} : \pi \in {}^{\text{H}}\text{hmn}_@ . \exists w' \in M_w [w' \in \pi(\lambda x_e . O_C \phi_x)]$ (Higher-order MS)

Assume that the modal base M maps the utterance world w to $\{w_1, w_2\}$, and that the committee is chaired by Andy alone in w_1 and by Billy alone in w_2 . This scenario is formally described as in (114).

⁴⁰Here RelExh is defined based on short answers, not sentential answers, because the interpretations of sentential answers vary as a function of the modal base. Consider the definition in (i), schematized in parallel to (112). Here Q^M stands for the Hamblin set of Q given the modal base M .

$$(i) \quad \forall M' [|M'_w| = 1 \wedge M'_w \subseteq M_w \wedge \exists p [w \in p \in Q^M \wedge w \in p \in Q^{M'}] \rightarrow \mathbf{DEP}(w, M', Q)]$$

Definition (i) is problematic because the two underlined parts are conflicting. For example, assume that the non-modalized question $? \phi_x$ has a true sentential answer ϕ_a in w_1 , and let $M_w = \{w_1, w_2\}$ and $M'_w = \{w_1\}$. Then, relative to both M and M' , $\diamond \phi_a$ is a true answer to the modalized question $? \diamond \phi_x$ in w ; however, $\llbracket \diamond \phi_a \rrbracket^M \neq \llbracket \diamond \phi_a \rrbracket^{M'}$, which causes a conflict. In contrast, definition (112) avoids this conflict because the interpretation of a short answer is independent of the modal base: for example, $\llbracket a \rrbracket^M = \llbracket a \rrbracket^{M'}$.

An anonymous reviewer of *NLS* points out that definition (112) of RelExh requires access to the modal base M , which is difficult to obtain from a compositional perspective. In the Kratzerian theory of modality, the modal base M is an argument variable of the modal verb. It is unclear how M can be retrieved when evaluating the interpretation of a modalized question as a whole, especially if the modal verb is embedded under another scopal expression (e.g., the anti-exhaustification operator *dou*). In Appendix B, I will explore a 'variable-free' analysis of modal bases, which defines modalized sentences as functions from modal bases to propositions. This analysis overcomes the technical difficulties in retrieving the modal base. It also allows us to define RelExh based on sentential answers.

$$(114) \text{ Let chair-the-committee} = \left[\begin{array}{l} w_1 \rightarrow \{a\} \\ w_2 \rightarrow \{b\} \\ \dots \end{array} \right] \text{ and } \left(\begin{array}{l} M_w = \{w_1, w_2\} \\ M_w^1 = \{w_1\} \\ M_w^2 = \{w_2\} \end{array} \right)$$

Given the modal base M , the first-order MS interpretation (113a) yields two true answers in w , namely, $\diamond O_C \phi_a$ and $\diamond O_C \phi_b$, verified by $\{w_1\}$ and $\{w_2\}$, respectively. This interpretation doesn't yield an exhaustive true answer and thus violates Dayal's exhaustivity presupposition. However, if this interpretation is evaluated relative to M^1 , which maps w to the singleton set $\{w_1\}$, $\diamond O_C \phi_a$ would be the unique exhaustive true answer in w ; likewise, if this interpretation is evaluated relative to M^2 , which maps w to $\{w_2\}$, $\diamond O_C \phi_b$ would be the unique exhaustive true answer in w . Hence, given the modal base M , the first-order MS interpretation (113a) satisfies RelExh in w .

(115) For the first-order MS interpretation (113a), we have:

$$\begin{array}{lll} \llbracket Q_{\text{FO}} \rrbracket_w^M = \{a, b\} & \text{ETA: non-existent} & \\ \llbracket Q_{\text{FO}} \rrbracket_w^{M^1} = \{a\} & \text{ETA: } \diamond O_C \phi_a & \text{(ETA abbreviates 'exhaustive true answer')} \\ \llbracket Q_{\text{FO}} \rrbracket_w^{M^2} = \{b\} & \text{ETA: } \diamond O_C \phi_b & \end{array}$$

This analysis also applies to the higher-order MS interpretation (113b). As seen below, the only change is that the higher-order interpretation allows for one more true answer, which is formed based on the Boolean disjunction $a^\uparrow \cup b^\uparrow$. However, this answer doesn't affect exhaustivity, since it remains partial regardless of the modal base.

(116) For the higher-order MS interpretation (113b), we have:

$$\begin{array}{lll} \llbracket Q_{\text{HO}} \rrbracket_w^M = \{a^\uparrow, b^\uparrow, a^\uparrow \cup b^\uparrow\} & \text{ETA: non-existent} & \\ \llbracket Q_{\text{HO}} \rrbracket_w^{M^1} = \{a^\uparrow, a^\uparrow \cup b^\uparrow\} & \text{ETA: } \diamond O_C \phi_a & \\ \llbracket Q_{\text{HO}} \rrbracket_w^{M^2} = \{b^\uparrow, a^\uparrow \cup b^\uparrow\} & \text{ETA: } \diamond O_C \phi_b & \end{array}$$

The relation between the two exhaustivity conditions exemplified above is generalized as follows:

(117) *Generalization of RelExh to MS interpretations*

The MS interpretations of $\ulcorner \text{Wh-A can } P? \urcorner$ satisfy RelExh if and only if the interpretations of the non-modalized question $\ulcorner \text{Wh-A } P? \urcorner$ satisfy Dayal's exhaustivity presupposition in every accessible world where P holds for an element of A .

I shall now revise the definitions of the answerhood operators as follows. The RelExh presupposition and the max-informativity condition are abbreviated as $\mathbf{REP}(w, M, \llbracket Q \rrbracket)$ and $\mathbf{MaxI}(\alpha, w, M, \llbracket Q \rrbracket)$, respectively. Applying an answerhood operator returns a set of max-informative true answers and triggers a RelExh presupposition.⁴¹

$$(118) \mathbf{MaxI}(\alpha, w, M, \llbracket Q \rrbracket) = 1 \text{ if and only if } \alpha \in \llbracket Q \rrbracket_w^M \text{ and } \forall \beta \in \llbracket Q \rrbracket_w^M [\llbracket Q \rrbracket^M(\beta) \not\subseteq \llbracket Q \rrbracket^M(\alpha)].$$

(119) Answerhood operators (modified from (40))

a. For complete true short answers:

$$\text{ANS}^S(w)(M)(\llbracket Q \rrbracket^M) = \mathbf{REP}(w, M, \llbracket Q \rrbracket) \cdot \{\alpha \mid \mathbf{MaxI}(\alpha, w, M, \llbracket Q \rrbracket)\}$$

⁴¹For readers who are familiar with Uegaki 2018, 2021: to account for the projection of the uniqueness presupposition in embeddings, it's better to assume that the RelExh presupposition is carried by each max-informative true answer, as opposed to by the answer set as a whole (see also fn. 36). With this assumption, the answerhood operators are defined as follows:

- (i) a. $\text{ANS}^S(w)(M)(\llbracket Q \rrbracket^M) = \{\mathbf{REP}(w, M, \llbracket Q \rrbracket) \cdot \alpha \mid \mathbf{MaxI}(\alpha, w, M, \llbracket Q \rrbracket)\}$
- b. $\text{ANS}^P(w)(M)(\llbracket Q \rrbracket^M) = \{\mathbf{REP}(w, M, \llbracket Q \rrbracket) \cdot \llbracket Q \rrbracket^M(\alpha) \mid \mathbf{MaxI}(\alpha, w, M, \llbracket Q \rrbracket)\}$

- b. For complete true propositional answers:
 $\text{ANS}^P(w)(M)(\llbracket Q \rrbracket^M) = \text{REP}(w, M, \llbracket Q \rrbracket) \cdot \{ \llbracket Q \rrbracket^M(\alpha) \mid \text{MaxI}(\alpha, w, M, \llbracket Q \rrbracket) \}$

6.2.2. Avoiding over-predictions of MS

As first raised in Sect. 4.2.3 and reiterated in Sect. 5.1.2, for the questions in (120) with a non-distributive predicate, the answer space derived in a first-order interpretation is not closed under conjunction. In such cases, without further constraints, applying Fox’s concept of answerhood would over-predict a MS interpretation.

- (120) a. Which children formed a team?
 b. Which children must form a team?

The RelExh presupposition solves the MS over-generation problem: in a multiple-team scenario, the first-order interpretations of (120a,b) violate RelExh, just as they violate Dayal’s exhaustivity presupposition.

Moreover, RelExh explains why existential indefinites do not license MS: RelExh allows the evaluation of exhaustivity to be relativized w.r.t. a smaller modal base, but not w.r.t. a smaller discourse domain.⁴² Hence, RelExh makes the same prediction as Dayal’s exhaustivity presupposition for non-modalized questions such as (84), *Which movie or movies did one of the boys watch?*

6.3. Predictions of RelExh regarding uniqueness

This subsection will explain how RelExh accounts for the observed uniqueness effects in questions. I will consider three types of questions with distinct modal force: (i) non-modalized questions (Sect. 6.3.1), (ii) *should*-questions (Sect. 6.3.2), and (iii) *can*-questions, interpreted as either MS (Sect. 6.3.3) or disjunctive MA (Sect. 6.3.4). Each type of question has multiple interpretations. In particular, for singular *wh*-questions, the higher-order interpretations differ from the first-order interpretations in that they allow for answers built out of Boolean disjunctions, which may affect the evaluation of exhaustivity.

In order to see whether the presented analysis completely accounts for uniqueness phenomena, we need to systematically check the question interpretations of singular *wh*-questions permitted by this analysis. In particular, we need to ensure that none of the permitted interpretations allows for the violation of uniqueness. We also need to ensure that every observed uniqueness inference, whether global or local, and whether universal or seemingly existential, can be generated from the permitted question interpretations. The following exemplifies the uniqueness inferences of modalized singular *wh*-questions and lists the relevant types of question interpretations:⁴³

⁴²‘Discourse domain’ refers to the interpretation parameter in model-theoretic semantics that determines the set of relevant individuals/entities. One might wonder why exhaustivity can be evaluated w.r.t. a smaller modal base but not w.r.t. a smaller discourse domain. As argued in Sect. 6.1, the essence of RelExh is that language users naturally shift the origo of interpretation to worlds where the state under discussion actually emerges. Evaluating exhaustivity w.r.t. singleton modal bases is just an analytical move to capture the perspectival shift. In contrast, discourse domain is independent of perspective.

Relatedly, as in situation semantics (Kratzer 1989), we may characterize propositions in terms of situations; thus we also have the option of evaluating exhaustivity w.r.t. minimal situations. Given this option, the RelExh-based analysis also applies to non-modalized questions like (19b) *Who has got a light?*, which share many MS-related properties with *can*-questions (van Rooij 2004). For this question, MS is permitted if exhaustivity can be evaluated w.r.t. minimal situations that exemplify the proposition *Someone has got a light*, which are situations consisting of a person, a light, and the possessing relation between the person and the light.

⁴³For example in (121b), ‘first-order (G)’ means that the first-order interpretation of (121a) yields global (G) uniqueness, and ‘narrow-scope higher-order (L)’ means that the narrow-scope higher-order interpretation of (121a) yields local (L) uniqueness.

- (121) Singular *have to*-questions
 - a. Example: “Which chapter do we have to assign to the students?”
 - i. Global uniqueness: *There is only one chapter that we have to assign.*
 - ii. Local uniqueness: *We have to assign at most one chapter.*
 - b. First-order (G); wide-scope higher-order (G); narrow-scope higher-order (L)
- (122) Singular *can*-questions without local exhaustification
 - a. Example: “Which chapter can we assign to the students?”
 - i. Global uniqueness: *There is only one chapter that we can assign.*
 - ii. Universal local uniqueness: *We can assign one chapter, but not more.*
 - b. First-order/higher-order MS (L); wide/narrow-scope disjunctive MA (L)
- (123) Singular *can*-questions with local exhaustification
 - a. Example: “Which SINGLE chapter can we assign to the students?”
 - i. Global uniqueness: *There is only one chapter that we can assign.*
 - ii. ‘Existential’ local uniqueness: *We can assign exactly one chapter to the students singly, aside from possibly available options of assigning more than one chapter simultaneously.*
 - b. First-order/higher-order MS (L); wide/narrow-scope disjunctive MA (L)

6.3.1. Uniqueness effects in non-modalized questions

The selection of modal base does not affect the interpretation of a non-modalized question. Therefore, RelExh carries forward the merits of Dayal’s exhaustivity presupposition in explaining the uniqueness effects in (82), repeated below:

- | | |
|---------------------------------------|--|
| (82) Which child came? | (Singular <i>wh</i> -question) |
| Which two children formed a team? | (Numeral-modified <i>wh</i> -question) |
| Did you invite Andy, Billy, or Cindy? | (As an alternative question) |

Moreover, the RelExh-based analysis of uniqueness extends to cases that are challenging for the presuppositional-*which* account of Hirsch and Schwarz (2020). RelExh can account for the uniqueness effect of (108b) and allows for the point-wise uniqueness effect of (109).

- | | |
|---|--|
| (108b) Which two or three students solved a problem together? | (Numeral-modified <i>wh</i> -question with a non-distributive predicate) |
| (109) Which boy watched which movie? | (Pair-list multiple- <i>wh</i> question) |

6.3.2. Uniqueness effects in *have to*-questions

For *have to*-questions (or any \square -questions, such as *must/should*-questions)⁴⁴ with a first-order interpretation, RelExh yields the same prediction as Dayal’s exhaustivity presupposition. For example in (124), in the first-order interpretation, the question calls for an answer naming an atomic chapter and presupposes that there is only one chapter that we have to assign to the students. Dayal’s exhaustivity presupposition is violated if there are multiple chapters that are assigned to the students in every accessible world. Since the uniqueness requirement is not satisfied in any accessible worlds,

⁴⁴The discussion here is centered on *have to*-questions just because the modal force of *have to* is unambiguously universal. However, the analysis and conclusions apply to any \square -questions.

the violation of Dayal’s exhaustivity presupposition cannot be salvaged by evaluating exhaustivity relative to a smaller modal base. Hence, RelExh yields global uniqueness. The same applies if this question has a wide-scope higher-order interpretation.

- (124) Q: Which chapter do we have to assign to the students?
 \rightsquigarrow *There is a unique chapter that we have to assign to the students.* (Global uniqueness)

A: Chapter 1.

However, when the same singular *have to*-question has a narrow-scope higher-order interpretation, RelExh and Dayal’s exhaustivity presupposition make different predictions. In (125), the question is exhaustively addressed by a narrow-scope disjunction, read as ‘There is no particular chapter that we have to assign to the students — we just need to chose between chapter 1 and chapter 2.’ (Spector 2007, 2008; Xiang 2021b; see Sect. 4.1.3). In this interpretation, the question implies local uniqueness.

- (125) Q: Which chapter do we have to assign to the students? (= (89))
 \rightsquigarrow *We shouldn’t assign more than one chapter to the students.* (Local uniqueness)

A: Chapter 1 or chapter 2, either is good. ($\Box(\phi_{c_1} \vee \phi_{c_2})$)

Given the modal base M specified in (126) (same as in (90b)), local uniqueness is satisfied in w' but not in w — w has an accessible world w_3 where uniqueness is violated. As argued in Sect. 5.1.3, Dayal’s exhaustivity presupposition cannot explain local uniqueness: given the modal base M , the *have to*-question in (125) has an exhaustive true answer in w (viz., $\Box(\phi_{c_1} \vee \phi_{c_2})$) despite the violation of uniqueness in w_3 .

$$(126) \text{ assign} = \left[\begin{array}{l} w_1 \rightarrow \{c_1\}, w_2 \rightarrow \{c_2\}, \\ w_3 \rightarrow \{c_1, c_2\}, \dots \end{array} \right], \quad M = \left[\begin{array}{l} w \rightarrow \{w_1, w_2, w_3\} \text{ (with uniq.-violation)} \\ w' \rightarrow \{w_1, w_2\} \text{ (without uniq.-violation)} \end{array} \right]$$

In contrast, the RelExh presupposition is satisfied in w' but violated in w . The true answer $\Box(\phi_{c_1} \vee \phi_{c_2})$ can be verified by three singleton sets of accessible worlds, namely $\{w_1\}, \{w_2\}, \{w_3\}$. RelExh is satisfied in w if and only if this question has an exhaustive true answer in w when interpreted relative to any modal base M' such that $M'_w = \{w_1\} / \{w_2\} / \{w_3\}$. With the assumed M , this requirement cannot be satisfied in w : given any M' such that $M'_w = \{w_3\}$, question (125) has two true answers in w (viz., $\Box\phi_{c_1}$ and $\Box\phi_{c_2}$) but no exhaustive true answer in w .

In sum, for a singular \Box -question, both RelExh and Dayal’s exhaustivity presupposition can predict a global-uniqueness inference, derived in the first-order interpretation or the wide-scope higher-order interpretation. However, only RelExh can predict a local-uniqueness effect, derived in a narrow-scope higher-order interpretation. Not only does this distinction argue that RelExh has advantages in accounting for local uniqueness; it also argues that RelExh is a mandatory condition for question interpretation, rather than merely a salvaging strategy for Dayal’s exhaustivity presupposition. If RelExh were optional, or if it only came into play in cases where Dayal’s exhaustivity presupposition is violated, we would expect local uniqueness to be optional or absent in singular \Box -questions.

6.3.3. Uniqueness effects in *can*-questions with a MS interpretation

According to the generalization in (117), the MS interpretations of the *can*-question (127b) satisfy RelExh if and only if the non-modalized question (127a) has a unique true answer in every accessible world where we assign any chapter(s) to the students. This condition is the desired universal local uniqueness inference.

- (127) a. Which chapter did we assign to the students?
 \rightsquigarrow *We assigned exactly one chapter to the students.* (Uniqueness)
- b. Which chapter can we assign to the students?
 \rightsquigarrow *We can assign exactly one chapter to the students, but not more.*
(Universal local uniqueness)

For a concrete illustration, let's interpret (127b) relative to the modal base M assumed in (126) above. In both w and w' , (127b) has the same set of true MS answers $\{\diamond\phi_{c_1}, \diamond\phi_{c_2}\}$. The MS interpretations of (127b) violate RelExh in w : since $\diamond\phi_{c_2}$ is a true answer in w and can be verified by $\{w_3\}$, RelExh requires that the question has an exhaustive true answer if interpreted relative to a modal base M' such that $M'_w = \{w_3\}$; however, when interpreted relative to such an M' , this question has two true answers in w (viz., $\diamond\phi_{c_1}$ and $\diamond\phi_{c_2}$) but no exhaustive true answer in w . In contrast, the MS interpretations of (127b) satisfy RelExh in w' since w' has no uniqueness-violating accessible world like w_3 .

In (103b), repeated as (128), we saw that the local-uniqueness inference appears to be existential when the *wh*-complement is modified by *SING*le: the question requests the addressee to name one chapter which can be the unique chapter that we assign to the students, but it does not rule out the possibility that a set containing more than one chapter could be assigned.

- (128) Which *SING*le chapter can we assign to the students?
 \rightsquigarrow *We can assign exactly one chapter to the students singly, aside from possibly available options of assigning more than one chapter simultaneously.* ('Existential' local uniqueness)

Why does the local-uniqueness inference in (103b)/(128) appear to be existential? Due to the modifier *SING*le, the question nucleus is parsed with local exhaustification, read as: 'Which chapter x is such that we can assign *only* x to the students?' Given the modal base M described in (126), the set of true MS answers to this question in w is $\{\diamond O\phi_{c_1}, \diamond O\phi_{c_2}\}$. The RelExh presupposition predicts the following condition: the non-modalized exhaustified question *Which chapter x is such that we assign only x to the students* has a unique true answer in every accessible world where there is a chapter x such that we only assign x to the students. This condition is fairly weak, since it only considers the accessible worlds where uniqueness is satisfied. For example, since none of the locally exhaustified true answers (e.g., $\diamond O_C\phi_{c_2}$) can be verified by $\{w_3\}$, the violation of uniqueness in w_3 doesn't affect RelExh. In sum, in a singular *can*-question, the local-uniqueness inference appears existential if the question nucleus is parsed with local exhaustification.

6.3.4. Uniqueness effects in *can*-questions with a disjunctive MA interpretation

Singular *can*-questions admit only disjunctive MA, not conjunctive MA. In (129), the MA answer can be expressed by an elided disjunction, as shown in (129a), but not by an elided conjunction or a plurality, as seen in (129b,c). This contrast argues that singular *which*-phrases may quantify over Boolean disjunctions, but not Boolean conjunctions or pluralities (Xiang 2021b; see also Sect. 5.1.1).

- (129) Which chapter can we assign to the students?
 \rightsquigarrow *We can assign exactly one chapter to the students, but not more.* (Universal local uniqueness)
- a. Chapter 1 or chapter 2.
b. # Chapter 1 and chapter 2.
c. # Chapters 1 and 2.

Incorporating this finding into the proposed derivation of disjunctive MA (Sect. 4.4.4), I argue that a singular *can*-question has a MA interpretation only if an anti-exhaustification operator (viz., *DOU*) appears above *can* and is associated with the higher-order *wh*-trace. Whether this *wh*-trace takes scope below or above *can* does not matter.

Let's return to local uniqueness. Section 6.3.3 has argued that RelExh can explain the universal local-uniqueness effect of the singular *can*-question in (129) if this question has a MS interpretation (see (127b)). Does the RelExh-based analysis apply if this question has a disjunctive MA interpretation? As I will argue below, in contrast to the case where this question has a MS interpretation, the RelExh presupposition alone cannot predict the observed universal local-uniqueness inference when this question has a disjunctive MA interpretation. Briefly, in a disjunctive MA interpretation, the violation of local uniqueness does not lead to a violation of RelExh.

Consider the following local-uniqueness-violating context, which is the same as the one we used for the MS interpretations. In this context, there are options to assign chapter 1 or chapter 2 by itself as well as the option to assign the two chapters together.

$$(130) \text{ assign} = \left[\begin{array}{l} w_1 \rightarrow \{c_1\}, w_2 \rightarrow \{c_2\} \\ w_3 \rightarrow \{c_1, c_2\}, \dots \end{array} \right], M_w = \{w_1, w_2, w_3\}$$

The answer spaces derived by disjunctive MA interpretations are as in Figures 8a,b, which differ only with respect to the scope of the higher-order *wh*-trace relative to the modal verb *can*. In both answer spaces, the disjunctive answer at the bottom is the unique exhaustive true answer.

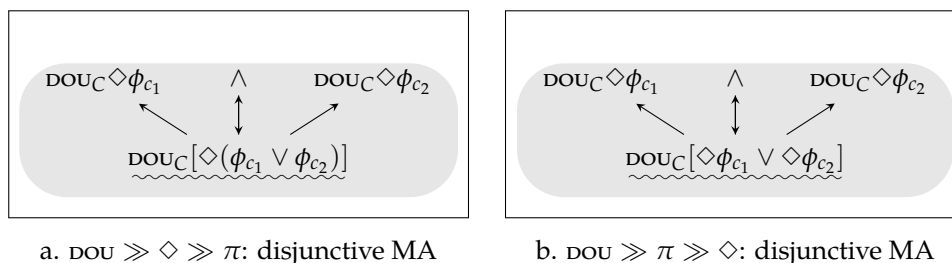


Figure 8: Answer space of (129) in a disjunctive MA interpretation

Now consider the predictions of RelExh. The true answer $\text{DOU}_C \diamond \phi_{c_2}$ (logically equivalent to $\diamond \phi_{c_2}$) can be verified by $\{w_3\}$; thus, RelExh requires that the question has an exhaustive true answer when it is evaluated relative to any modal base M'_w such that $M'_w = \{w_3\}$. Sect. 6.3.3 has argued that the MS interpretations of this question do not satisfy this requirement in the given w : these MS interpretations do not satisfy exhaustivity when evaluated relative to an M'_w such that $M'_w = \{w_3\}$. However, here the answer spaces derived in disjunctive MA interpretations include also a FC-disjunctive answer, $\text{DOU}_C [\diamond \phi_{c_1} \vee \diamond \phi_{c_2}]$ or $\text{DOU}_C [\diamond (\phi_{c_1} \vee \phi_{c_2})]$, both of which are logically equivalent to $\diamond \phi_{c_1} \wedge \diamond \phi_{c_2}$. Given any M'_w such that $M'_w = \{w_3\}$, this FC-disjunctive answer is an exhaustive true answer to this *can*-question in w , and thus RelExh is not violated in w . Hence, RelExh alone cannot explain the observed universal local-uniqueness effect if (129) has a disjunctive MA interpretation.

How else then does my account avoid under-generating uniqueness? I'd like to suggest that, before RelExh applies, the FC-disjunctive answers $\text{DOU}_C [\diamond \phi_{c_1} \vee \diamond \phi_{c_2}]$ and $\text{DOU}_C [\diamond (\phi_{c_1} \vee \phi_{c_2})]$ have been ruled out due to violations of Relativized Exclusivity (RelExcl), a condition that has been independently motivated to account for the modal obviation effect in the licensing of universal FC. As proposed in Sect. 4.4.3, the anti-exhaustification operator *DOU* carries a RelExcl presupposition: for

every anti-excludable alternative ϕ stronger than the prejacent, every minimal subset of the accessible worlds that verifies ϕ also verifies the exhaustification of ϕ (for formal definitions, see (66) and (141)). When DOU applies to a \diamond -disjunction, no matter whether the disjunction takes scope below or above the existential modal, the RelExcl presupposition yields a definedness condition as follows:

- (131) DOU_C [John $\text{or}_{[+D]}$ Mary can teach Intro Chinese]
- a. Without local exhaustification: $\text{DOU}_C[\diamond(\phi_j \vee \phi_m)]$ and $\text{DOU}_C[\diamond\phi_j \vee \diamond\phi_m]$ are defined only if $\diamond O\phi_j \wedge \diamond O\phi_m \wedge \neg\diamond(\phi_j \wedge \phi_m)$.
 - b. With local exhaustification: $\text{DOU}_C[\diamond(O\phi_j \vee O\phi_m)]$ and $\text{DOU}_C[\diamond O\phi_j \vee \diamond O\phi_m]$ are defined only if $\diamond O\phi_j \wedge \diamond O\phi_m$.

As seen in (131a), if the prejacent sentence is parsed without local exhaustification, RelExcl yields the condition that the two distinct disjuncts cannot be simultaneously true. For the same reason, for the singular *can*-question (129), if $\diamond(\phi_{c_1} \wedge \phi_{c_2})$ is true, the disjunctive answers $\text{DOU}_C[\diamond(\phi_{c_1} \vee \phi_{c_2})]$ and $\text{DOU}_C[\diamond\phi_{c_1} \vee \diamond\phi_{c_2}]$ are undefined. Once the disjunctive answers that violate RelExcl are removed from the answer space, the RelExh presupposition predicts a universal local-uniqueness effect for the disjunctive MA interpretations in the same way as for the MS interpretations (see (127b)).

This analysis also applies to (132) below, repeated from (128). As argued in Sect. 6.3.3, due to the modifier *SINgle*, the nucleus of this question is parsed with local exhaustification, read as: ‘Which chapter x is such that we can assign *only* x to the students?’ As in the case of (131b), when parsed with local exhaustification, the disjunctive answers $\text{DOU}_C[\diamond(O\phi_{c_1} \vee O\phi_{c_2})]$ and $\text{DOU}_C[\diamond O\phi_{c_1} \vee \diamond O\phi_{c_2}]$ satisfy RelExcl as long as $\diamond O\phi_{c_1}$ and $\diamond O\phi_{c_2}$ are true, regardless of whether $\diamond(\phi_{c_1} \wedge \phi_{c_2})$ is false or true. Hence for (132), RelExh predicts a seemingly existential local-uniqueness effect for the disjunctive MA interpretations in the same way as it does for the MS interpretations (see (128)).

- (132) Q: Which *SINgle* chapter can we assign to the students?
 \rightsquigarrow We can assign exactly one chapter to the students singly, aside from possibly available options of assigning more than one chapter simultaneously. (‘Existential’ local uniqueness)
- A: Chapter 1 or chapter 2.

In sum, singular *can*-questions admit FC-disjunctive answers built out of Boolean disjunctions. When such answers are available, RelExh alone cannot explain the universal local-uniqueness effects. I argue that whether a FC-disjunctive answer is available is independently restricted by RelExcl , a presupposition carried by the anti-exhaustification operator. Once the disjunctive answers that violate RelExcl are removed, the RelExh presupposition can account for the universal local-uniqueness effects.

6.4. Interim summary

To predict the distribution of MS interpretations and uniqueness effects in questions, this section has assumed a ‘Relativized Exhaustivity (RelExh)’ presupposition as a mandatory condition for question semantics that reflects the perspective shift to worlds where the state under discussion actually emerges. Accordingly, for modalized questions, semantic properties such as exhaustivity are evaluated relative to the accessible worlds, not the utterance world. As I hope to have shown, within the framework of model-theoretic compositional semantics, this intrinsic propensity of questions to be perspectival can be formally captured by requiring that Dayal’s exhaustivity presupposition must be satisfied relative to every modal base that introduces a singleton set of accessible worlds which verifies a true answer.

The RelExh presupposition has many advantages: it allows for MS interpretations, carries forward the merits of Dayal’s exhaustivity presupposition in deriving uniqueness effects, and avoids over-generating MS interpretations for non-*can*-questions. Moreover, in contrast to Dayal’s exhaustivity presupposition, RelExh can account for the local-uniqueness effects of modalized questions. The predictions of Dayal’s exhaustivity presupposition and RelExh are summarized as follows:

Modal type	Reading type	Dayal’s E.P.	RelExh
No modal	\pm uniq.	✓	✓
□-modal	– uniq.	✓	✓
	global uniq.	✓	✓
	local uniq.	✗	✓
◇-modal	– uniq. MS	✗	✓
	– uniq. MA	✓	✓
	global uniq.	✓	✗
	local uniq. MS	✗	✓
	local uniq. MA	✗	✓

Table 3: Predictions of Dayal’s exhaustivity presupposition and the RelExh presupposition

Note that, in the case of a ◇-question, RelExh predicts only a local-uniqueness effect. To allow for interpretations with global uniqueness, we can either assume that Dayal’s exhaustivity presupposition is applied optionally, or assume that questions have an optional uniqueness condition, namely, that there can be only one complete true answer.

7. Conclusions

This paper made three contributions to the study of MS interpretations and uniqueness effects in questions. First, observing that MS answers are subject to a ‘mention-one-only’ requirement which cannot be explained by pragmatic factors, I argued that MS interpretations are primarily licensed by grammatical factors, especially the presence of the modal verb *can*. Further, given that this modal verb has to be interpreted within the question nucleus, I argued that the MS/MA ambiguity in *can*-questions should be analyzed in terms of structural ambiguities within the question nucleus. For cases where MS is licensed by a non-exhaustive conversational goal, I argued that the context may provide covert restrictions to the question nucleus and turn the question into a *can*-question.

Second, taking insights from Fox 2013, I derived MS and MA interpretations of *can*-questions with a single non-exhaustive max-informativity-based definition of answerhood and attributed the MS/MA contrast to structural variations within the question nucleus. I argued that MA interpretations arise if one of the following conditions is met, and MS interpretations arise otherwise: (i) the higher-order *wh*-trace takes scope above the modal *can*, or (ii) an anti-exhaustification operator *DOU* (\approx the Mandarin FC-triggering particle *dou*) appears above *can* and is associated with the higher-order *wh*-trace. In particular, condition (i) yields interpretations calling for a conjunctive MA answer, and condition (ii) yields an interpretation calling for a disjunctive MA answer. With respect to disjunctive MA, I also proposed a ‘Relativized Exclusivity’ condition (cf. Dayal’s (2013) Viability constraint). This condition uniformly accounts for the modal obviation effect in licensing universal FC and the distribution of FC-disjunctive answers.

However, allowing non-exhaustive answers to be complete would cause a troubling conflict with ‘Dayal’s exhaustivity presupposition’, which says that a question must have an exhaustive true answer.

This condition is crucial in accounting for the uniqueness effects in questions. Hence, last and most importantly, I proposed that question interpretations can violate Dayal’s exhaustivity presupposition but mandatorily presuppose ‘Relativized Exhaustivity’. This condition solves the dilemma between uniqueness and MS, avoids over-generating MS interpretations for non-*can*-questions, and is advantageous in deriving local-uniqueness effects in modalized questions.

A. Choice questions versus MS questions

As seen in (84), repeated in (133), *wh*-questions with an existential indefinite are ambiguous between a choice interpretation and an individual interpretation. In contrast to the individual interpretation, which has been discussed in Sect. 5.1.2, the choice interpretation is globally non-exhaustive: as seen in (133a), a choice answer only needs to specify one boy-movie(s) pair.

(133) (Among the boys under consideration, Andy watched *Ironman* and *Spiderman*, and Billy watched only *Hulk*. Clark didn’t watch any movies.)

Which movie or movies did one of the boys watch?

- a. ‘Name any/one boy *x*, and then tell me: Which movie(s) did *x* watch?’ (Choice)
 - i. Andy watched *Ironman* and *Spiderman*.
 - ii. Billy watched *Hulk*.
- b. ‘Which movie(s) *y* is/are such that one of the boys watched *y*?’ (Individual)
 - i. One (of the boys) watched *Ironman* and *Spiderman*, and one watched *Hulk*.
 - ii. #One (of the boys) watched *Hulk*.

Should the non-exhaustive choice interpretation be treated as a variant of the MS interpretation? Dayal (2017: Sect. 3.2.1) has discussed two reasons for keeping choice interpretations separate from MS interpretations. One reason draws on the distinction w.r.t. the morphological form of the *wh*-expression in these questions, and the other is concerned with the availability of MS interpretations in questions with a universal quantifier. She calls questions with a choice interpretation ‘choice questions’, in contrast to ‘MS questions’. This appendix presents three additional arguments.

First, choice questions admit only ‘choose-some’ interpretations, not ‘choose-all’ interpretations. For example, while (134a) admits a MA interpretation, (134b) can never be read as the consultant knowing every store’s opening time. To convey this meaning, one has to replace the existential indefinite with a universal quantifier, as in (134c). This contrast argues that the MS/MA ambiguity in *can*-questions and the choose-some/choose-all distinction in questions with quantifiers have different origins.

- (134) a. The consultant knows where we can get coffee.
 b. The consultant knows when one of the stores opens.
 c. The consultant knows when each of the stores open.

Second, disjunctive answers to choice questions do not have a FC reading. In (135), the disjunctive answer only has an epistemic ignorance reading. The unavailability of a FC reading argues that the machinery that makes a *can*-question congruent with FC-disjunctive answers is not available to choice questions. As argued in Sect. 4.4, the disjunctive MA interpretation of a *can*-question is derived by applying an anti-exhaustification operator to the local IP. Nevertheless, as commonly assumed, the existential indefinite in a choice question occupies a fairly high position at LF (Groenendijk and

Stokhof 1984; Chierchia 1993; Szabolcsi 1997b; Krifka 2001b; Dayal 2017; Xiang 2019, 2021c; a.o.), and therefore it cannot interact with an anti-exhaustification operator. Hence, it is expected by the presented proposal that choice questions do not have FC-disjunctive answers.

- (135) Q: ‘Which movie did one of the two boys watch?’ (Choice)
 A: ‘Andy watched *Hulk*, or Billy watched *Ironman*.’
 (Available: ‘Either Andy watched *Hulk* or Billy watched *Ironman*, I don’t know which.’)
 (Unavailable: ‘Andy watched *Hulk*, and Billy watched *Ironman*.’)

Third, the two types of questions behave quite differently w.r.t. local uniqueness. In a singular *can*-question, the local uniqueness presupposition is universal — it applies to every accessible world that verifies a true answer (see (103)–(106) and (127)/(129)). In choice questions, however, the local uniqueness presupposition is existential. For example, the choice question in (135) implies that one of the boys watched exactly one movie, not that each boy watched at most one movie. This contrast argues that the local-uniqueness interpretations have different derivations in these two types of questions. See Xiang 2019, 2021c for an analysis that derives the existential uniqueness effects of singular choice questions.

B. A variable-free treatment of modal bases

As mentioned, the definitions for Relativized Exclusivity (RelExcl) and Relativized Exhaustivity (RelExh) face some technical challenges.

Consider RelExcl first. As discussed in fn. 28, the definition for RelExcl in (66), repeated below, is syncategorematic: we cannot define this condition as part of the lexical meaning of *dou*/DOU. Moreover, we cannot precisely define the value of C' , which should vary with the value of M' .

- (136) *Relativized Exclusivity* (repeated from (66))
 $\llbracket dou_C(S) \rrbracket^M / \llbracket DOU_C(S) \rrbracket^M$ is defined in w only if
 $\forall \phi [\llbracket \phi \rrbracket^M \in \text{ANTIEXCL}(\llbracket S \rrbracket^M, C) \wedge \llbracket \phi \rrbracket^M \subset \llbracket S \rrbracket^M$
 $\rightarrow \forall M'_{(s,st)} [M'_w \text{ is a minimal subset of } M_w \text{ s.t. } \llbracket \phi \rrbracket^{M'}(w) = 1 \rightarrow O_{C'}(\llbracket \phi \rrbracket^{M'})(w) = 1]]$.

Moreover, in the Kratzerian theory of modality (Kratzer 1977, 1991), the modal base is treated as a free argument variable of the modal verb and is interpreted via an assignment function, as formalized in (137). With this treatment, it is unclear how the interpretation of a modalized sentence can make reference to the modal base, especially in cases where the modal verb is embedded (e.g., under *dou*).

- (137) a. $\llbracket can_M \phi \rrbracket^g = \lambda w. \exists w' \in g(M)(w) [\llbracket \phi \rrbracket^g(w) = 1]$
 b. $\llbracket should_M \phi \rrbracket^g = \lambda w. \forall w' \in g(M)(w) [\llbracket \phi \rrbracket^g(w) = 1]$

To address these technical problems, I here explore a ‘variable-free’ treatment of modal bases. This treatment is inspired by the variable-free analysis of pronouns (Jacobson 1999, 2014). In the Heim-and-Kratzer tradition, pronouns are translated as free variables and are interpreted via an assignment function (e.g., $\llbracket it_i \rrbracket^g = g(i)$). In contrast, the variable-free analysis of Jacobson defines a pronoun as an $\langle e, e \rangle$ -type identity function (e.g., $\llbracket it \rrbracket = \lambda x_e. x$). Also, an expression that contains a free pronoun has an e -type abstraction passed up from this pronoun (e.g., $\llbracket it \text{ arrived} \rrbracket = \lambda x_e. \text{arrive}(x)$). Extending this treatment to modal bases, I assume that modalized sentences denote functions from modal bases to propositions, as defined in (138).

- (138) a. $\llbracket can \phi \rrbracket = \lambda M_{(s,st)} \lambda w. \exists w' \in M_w [\llbracket \phi \rrbracket(w) = 1]$

$$\text{b. } \llbracket \textit{should } \phi \rrbracket = \lambda M_{\langle s, st \rangle} \lambda w. \forall w' \in M_w [\llbracket \phi \rrbracket (w) = 1]$$

Further, in analogy to the **g**-rule (shorthand for ‘Geach rule’) in Jacobson’s variable-free semantics, I assume the following type-shifting operation \mathcal{MB} which allows a sentential operator to apply to a modalized sentence (of type $\langle sst, st \rangle$) and passes up the abstraction of the modal base.

$$(139) \quad \text{For any sentential expression } F \text{ of type } \langle st, \sigma \rangle, \mathcal{MB}(F) \text{ is an expression of type } \langle \langle sst, st \rangle, \langle sst, \sigma \rangle \rangle \text{ such that } \llbracket \mathcal{MB}(F) \rrbracket = \lambda \alpha_{\langle sst, st \rangle} \lambda M_{\langle s, st \rangle} \cdot \llbracket F \rrbracket (\alpha(M)).$$

For example, the application of \mathcal{MB} shifts the assertion of DOU from (140a) to (140b). $\mathcal{MB}(\text{DOU}_C)$ takes a $\langle sst, st \rangle$ -type sentence as its argument and returns a $\langle sst, st \rangle$ -type sentence. The value of the variable C as well is shifted from a set of propositions into a set of functions of type $\langle sst, st \rangle$.

(140) Asserted meaning of DOU

$$\text{a. } \llbracket \text{DOU}_C \rrbracket = \lambda p_{\langle s, t \rangle} \lambda w. p(w) = 1 \wedge \forall q \in \text{ANTIEXCL}(p, C) [O_C^{\text{IE}}(q)(w) = 0] \quad (= (64))$$

$$\text{b. } \llbracket \mathcal{MB}(\text{DOU}_C) \rrbracket \\ = \lambda \theta_{\langle sst, st \rangle} \lambda M_{\langle s, st \rangle} \lambda w. \theta(M)(w) = 1 \wedge \forall q \in \text{ANTIEXCL}(\theta(M), C^M) [O_{C^M}^{\text{IE}}(q)(w) = 0], \\ \text{where for any } M_{\langle s, st \rangle}, C^M := \{\rho(M) \mid \rho_{\langle sst, st \rangle} \in C\}$$

The above assumptions allow us to redefine the RelExcl presupposition of DOU as in (141). To evaluate the exclusivity of a modalized sentence θ w.r.t. a modal base M' , all we need is to apply each of its modalized alternatives δ to M' .

(141) *Relativized Exclusivity* (modified from (136))

$$\text{For any } \theta \text{ of type } \langle sst, st \rangle \text{ and } M \text{ of type } \langle s, st \rangle, \llbracket \mathcal{MB}(\text{DOU}_C) \rrbracket (\theta)(M)(w) \text{ is defined only if} \\ \forall \delta_{\langle sst, st \rangle} [\delta(M) \in \text{ANTIEXCL}(\theta(M), C^M) \wedge \delta(M) \subset \theta(M) \\ \rightarrow \forall M' [M'_w \text{ is a minimal subset of } M_w \text{ s.t. } \delta(M')(w) = 1 \rightarrow O_{C^{M'}}^{\text{IE}}(\delta(M'))(w) = 1]].$$

Next, let’s see how the variable-free analysis of modal bases affects Relativized Exhaustivity (RelExh). As discussed in fn. 40, the definition of RelExh in (112) requires the extraction of short answers from the question denotation, because the semantics of a sentential answer varies with the modal base. Moreover, for the same reason mentioned in the discussion of RelExcl, this definition requires access to the value of M , which is difficult to obtain in a compositional perspective.

(142) *Relativized Exhaustivity* (repeated from (112))

$$\text{a. For any modal base } M, \text{ a question } Q \text{ is defined in } w \text{ only if} \\ \forall M'_{\langle s, st \rangle} [|M'_w| = 1 \wedge M'_w \subseteq M_w \wedge \exists \alpha [\alpha \in \llbracket Q \rrbracket_w^M \wedge \alpha \in \llbracket Q \rrbracket_w^{M'}] \rightarrow \text{DEP}(w, M', \llbracket Q \rrbracket)] \\ \text{b. } \text{DEP}(w, M', \llbracket Q \rrbracket) := \exists \alpha [\alpha \in \llbracket Q \rrbracket_w^M \wedge \forall \beta [\beta \in \llbracket Q \rrbracket_w^{M'} \rightarrow \llbracket Q \rrbracket^{M'}(\beta) \subseteq \llbracket Q \rrbracket^{M'}(\alpha)]]$$

However, with the variable-free analysis of modal bases, we can now define the root denotation of a modalized question as a Hamblin set Q , each member of which is a function from a modal base to a propositional answer. The definition of RelExh can be revised as follows, where Q (of type $\langle \langle sst, st \rangle, t \rangle$) stands for a Hamblin set and $Q^M := \{\rho(M) \mid \rho_{\langle sst, st \rangle} \in Q\}$.

(143) *Relativized Exhaustivity* (modified for Hamblin sets)

$$\text{a. For any modal base } M, \text{ a question } Q \text{ is defined in } w \text{ only if} \\ \forall M'_{\langle s, st \rangle} [|M'_w| = 1 \wedge M'_w \subseteq M_w \wedge \exists \theta [w \in \theta(M) \in Q^M \wedge w \in \theta(M') \in Q^{M'}] \rightarrow \text{DEP}(w, M', Q)] \\ \text{b. } \text{DEP}(w, M', Q) := \exists \theta [w \in \theta(M') \in Q^M \wedge \forall \delta [w \in \delta(M') \in Q^{M'} \rightarrow \theta(M') \subseteq \delta(M')]]$$

Acknowledgements [To be added ...]

References

- Abrusán, Márta. 2014. *Weak Island Semantics*. Oxford: Oxford University Press.
- Aloni, Maria. 2007. Free choice, modals, and imperatives. *Natural Language Semantics* 15:65–94.
- Bar-Lev, Moshe E, and Danny Fox. 2020. Free choice, simplification, and innocent inclusion. *Natural Language Semantics* 28:175–223.
- Barker, Chris. 2010. Free choice permission as resource-sensitive reasoning *Semantics and Pragmatics* 3, Article 10:1–38.
- Beck, Sigrid, and Hotze Rullmann. 1999. A flexible approach to exhaustivity in questions. *Natural Language Semantics* 7:249–298.
- Caponigro, Ivano, and Kathryn Davidson. 2011. Ask, and tell as well: Question–answer clauses in American Sign Language. *Natural Language Semantics* 19:323–371.
- Chierchia, Gennaro. 1993. Questions with quantifiers. *Natural Language Semantics* 1:181–234.
- Chierchia, Gennaro. 2006. Broaden your views: Implicatures of domain widening and the “logicality” of language. *Linguistic Inquiry* 37:535–590.
- Chierchia, Gennaro. 2013. *Logic in grammar: Polarity, free choice, and intervention*. Oxford: Oxford University Press.
- Chierchia, Gennaro, and Ivano Caponigro. 2013. Questions on questions and free relatives. Presentation at the 18th Sinn und Bedeutung, University of the Basque Country (UPV/EHU) in Vitoria-Gasteiz, September 2013. https://scholar.harvard.edu/files/chierchia/files/frs_and_qs_iii_2013-9-11.pdf.
- Chierchia, Gennaro, Danny Fox, and Benjamin Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In *An international handbook of natural language meaning*, ed., Claudia Maienborn, Klaus von Heusinger, and Paul Portner, 2297–2332. Berlin: Mouton de Gruyter.
- Dayal, Veneeta. 1996. *Locality in Wh Quantification: Questions and Relative Clauses in Hindi*. Dordrecht: Kluwer.
- Dayal, Veneeta. 2013. A viability constraint on alternatives for free choice. In *Alternatives in semantics*, ed., Anamaria Fălăuş, 88–122. London: Palgrave Macmillan.
- Dayal, Veneeta. 2017. *Questions*. Oxford: Oxford University Press.
- Dotlačil, Jakub, and Floris Roelofsen. 2019. Dynamic inquisitive semantics: Anaphora and questions. In *Proceedings of sinn und bedeutung 23*, ed., M.Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally, and Cristina Real-Puigdollers, 365–382. <https://doi.org/10.18148/sub/2019.v23i1.538>.
- Dotlačil, Jakub, and Floris Roelofsen. 2021. A dynamic semantics of single-*wh* and multiple-*wh* questions. In *Proceedings of Semantics and Linguistic Theory 30*, ed., Joseph Rhyne, Kaelyn Lamp, Nicole Dreier, and Chloe Kwon, 376–395. Washington, D.C.: LSA. DOI:<https://doi.org/10.3765/salt.v30i0.4839>.

- von Fintel, Kai, Danny Fox, and Sabine Iatridou. 2014. Definiteness as maximal informativeness. In *The art and craft of semantics: A festschrift for Irene Heim*, vol. 1, ed., Luka Crnic and Uli Sauerland, 165–174. MIT Working Papers in Linguistics.
- Fox, Danny. 2007. Free choice disjunction and the theory of scalar implicatures. In *Presupposition and implicature in compositional semantics*, ed., Uli Sauerland and Penka Stateva, 71–120. New York: Palgrave Macmillan.
- Fox, Danny. 2012. Multiple *wh*-questions: Uniqueness, pair-list and second order questions. MIT seminar handouts. <http://lingphil.mit.edu/papers/fox/uniqueness-pair-list.pdf>.
- Fox, Danny. 2013. Mention-some interpretations. MIT seminar handouts. <http://lingphil.mit.edu/papers/fox/class1-3.pdf>.
- Fox, Danny. 2018. Partition by exhaustification: Comments on Dayal 1996. In *Proceedings of Sinn und Bedeutung 22*, ed., Uli Sauerland and Stephanie Solt, 403–434. Berlin: Leibniz-Centre General Linguistics. <https://semanticsarchive.net/Archive/jQ5MTcyZ/>.
- Fox, Danny. 2020. Partition by exhaustification: Towards a solution to Gentile and Schwarz’s puzzle. Manuscript, MIT. <https://semanticsarchive.net/Archive/TljzGNjz/>.
- Fox, Danny, and Martin Hackl. 2007. The universal density of measurement. *Linguistics and Philosophy* 29:537–586.
- George, B. R. 2011. Question embedding and the semantics of answers. Ph.D. Dissertation, University of California, Los Angeles.
- Ginzburg, Jonathan. 1995. Resolving questions, i. *Linguistics and Philosophy* 18:459–527.
- Ginzburg, Jonathan, and Ivan Sag. 2000. *Interrogative investigations*. Stanford: CSLI Publications.
- Groenendijk, Jeroen, and Martin Stokhof. 1984. On the semantics of questions and the pragmatics of answers. *Varieties of formal semantics* 3:143–170.
- Hausser, Roland, and Dietmar Zaefferer. 1979. Questions and answers in a context-dependent Montague grammar. In *Formal semantics and pragmatics for natural languages*, ed., Franz Guenther and Siegfried J. Schmidt, 339–358. Dordrecht: Reidel.
- Hausser, Roland R. 1983. The syntax and semantics of English mood. In *Questions and answers*, ed., Ferenc Kiefer, 97–158. Dordrecht: Springer.
- Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9:183–221.
- Heim, Irene. 1994. Interrogative semantics and Karttunen’s semantics for *know*. In *Proceedings of the 9th Annual Conference of the Israel Association for Theoretical Linguistics (IATL) and of the Workshop on Discourse*, ed., Rhonna Buchalla and Anita Mittwoch, volume 1, 128–144. Jerusalem: IATL. <https://semanticsarchive.net/Archive/jUzYjkl0/Interrogative%2094.pdf>.
- Hirsch, Aron, and Bernhard Schwarz. 2020. Singular *which*, mention-some, and variable scope uniqueness. In *Proceedings of Semantics and Linguistic Theory 29*, ed., Katherine Blake, Forrest Davis, Kaelyn Lamp, and Joseph Rhyne, 748–767. Washington, D.C.: LSA. <https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/29.748>.
- Huang, C.-T. James. 1982. Move WH in a language without *wh* movement. *The Linguistic Review* 1:369–416.
- Jacobson, Pauline. 1999. Towards a variable-free semantics. *Linguistics and Philosophy* 22:117–185.

- Jacobson, Pauline. 2014. *Compositional semantics: An introduction to the syntax/semantics interface*. Oxford: Oxford University Press.
- Karttunen, Lauri. 1974. Presupposition and linguistic context. *Theoretical Linguistics* 1:181–194.
- Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1:3–44.
- Klinedinst, Nathan, and Daniel Rothschild. 2011. Exhaustivity in questions with non-factives. *Semantics and Pragmatics* 4:1–23.
- Kratzer, Angelika. 1977. What ‘must’ and ‘can’ must and can mean. *Linguistics and Philosophy* 1:337–355.
- Kratzer, Angelika. 1981. The notional category of modality. In *Words, worlds, and contexts: New approaches in word semantics*, ed., H. Eikmeyer and H. Rieser, 38–74. Berlin: Walter de Gruyter.
- Kratzer, Angelika. 1989. An investigation of the lumps of thought. *Linguistics and philosophy* 12:607–653.
- Kratzer, Angelika. 1991. Modality. In *Semantics: An international handbook of contemporary research*, ed., A. von Stechow and D. Wunderlich, 639–650. Berlin: de Gruyter.
- Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In *The Proceedings of the Third Tokyo Conference on Psycholinguistics*, ed., Yukio Otsu, 1–25. Tokyo: Hituzi Syobo. Reprinted in 2017, in *Contrastiveness in information structure, alternatives and scalar implicatures*, ed., Lee C., Kiefer F., Krifka M, 123–143. Springer, Cham.
- Krifka, Manfred. 2001a. For a structured meaning account of questions and answers. *Audiatur vox sapientiae. A festschrift for Arnim von Stechow*, ed., C. Féry and W. Sternefeld, 52:287–319. Berlin: Akademie Verlag.
- Krifka, Manfred. 2001b. Quantifying into question acts. *Natural Language Semantics* 9:1–40.
- Kobayashi, Filipe Hisao, and Vincent Rouillard. 2021. High and low uniqueness in singular *wh*-interrogatives. In *Proceedings of Semantics and Linguistic Theory 30*, ed., Joseph Rhyne, Kaelyn Lamp, Nicole Dreier, and Chloe Kwon, 714–733. Washington, D.C.: LSA <http://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/30.714>.
- Lahiri, Utpal. 2002. *Questions and answers in embedded contexts*. Oxford: Oxford University Press.
- Li, Haoze. 2019. Deriving short answers from Hamblin sets: A case study of Mandarin *wh*-conditionals. In *Proceedings of Semantics and Linguistic Theory 29*, ed., Katherine Blake, Forrest Davis, Kaelyn Lamp, and Joseph Rhyne, 138–159. Washington, D.C.: LSA.
- Li, Haoze. 2021. Mandarin *wh*-conditionals: A dynamic question approach. *Natural Language Semantics*. Online-first articles.
- Li, Jie. 1995. *Dou* and *wh*-questions in Mandarin Chinese. *Journal of East Asian Linguistics* 29:313–323.
- Li, Xiaoguang. 1997. Deriving distributivity in Mandarin Chinese. PhD Dissertation, University of California, Irvine.
- Lin, Jo-Wang. 1998. Distributivity in Chinese and its implications. *Natural Language Semantics* 6:201–243.
- Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, use, and interpretation of language*, ed., Christoph Schwarze Rainer Bäuerle and Arnim von Stechow, 302–323. Berlin: De Gruyter.
- Liu, Mingming. 2016a. Varieties of alternatives: Mandarin focus particles. *Linguistics and Philosophy* 40:61–95.

- Liu, Mingming. 2016b. Mandarin *wh*-conditionals as interrogative conditionals. In *Proceedings of Semantics and Linguistic Theory 26*, ed., Mary Moroney, Carol-Rose Little, Jacob Colvard, and Dan Burgdorf, 814–835. Ithaca, NY: CLC Publications. <https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/26.814>.
- Matthewson, Lisa. 2001. Quantification and the nature of crosslinguistic variation. *Natural Language Semantics* 9:145–189.
- Menéndez-Benito, Paula. 2010. On universal free choice items. *Natural Language Semantics* 18:33–64.
- Nicolae, Andreea C. 2013. *Any* questions? Polarity as a window into the structure of questions. Ph.D. Dissertation, Harvard University.
- Nicolae, Andreea C. 2015. Questions with NPIs. *Natural Language Semantics* 1:21–76.
- Portner, Paul. 2009. *Modality*. Oxford: Oxford University Press.
- van Rooij, Robert. 2003. Questioning to resolve decision problems. *Linguistics and Philosophy* 26:727–763.
- van Rooij, Robert. 2004. Utility of mention-some questions. *Research on Language and Computation* 2:401–416.
- van Rooij, Robert, and Katrin Schulz. 2004. Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information* 13:491–519.
- Rooth, Mats. 1996. Focus. In *The handbook of contemporary semantic theory*, ed., Shalom Lappin, 271–297. Oxford: Blackwell Publishers.
- Rullmann, Hotze, and Sigrid Beck. 1998. Presupposition projection and the interpretation of *which*-questions. In *Semantics and Linguistic Theory 8*, ed., Devon Strolovitch and Aaron Lawson, 215–232. <https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/2811>.
- Sauerland, Uli, Jan Anderssen, and Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In *Linguistic evidence: Empirical, theoretical, and computational perspectives*, ed., Stephan Kepser and Marga Reis, 413–434. Berlin: Mouton de Gruyter.
- Schulz, Katrin, and Robert van Rooij. 2006. Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy* 29:205–250.
- Sharvy, Richard. 1980. A more general theory of definite descriptions. *The Philosophical Review* 89:607–624.
- Simons, Mandy. 2005. Dividing things up: The semantics of *or* and the modal/*or* interaction. *Natural Language Semantics* 13(3): 271–316.
- Spector, Benjamin. 2007. Modalized questions and exhaustivity. In *Proceedings of Semantics and Linguistic Theory 17*, ed., M. Gibson and T. Friedman, 282–299. Ithaca, NY: CLC Publications. <https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/2962>.
- Spector, Benjamin. 2008. An unnoticed interpretation for *wh*-questions: Elided answers and weak islands. *Linguistic Inquiry* 39:677–686.
- Spector, Benjamin, and Marta Abrusán. 2011. A semantics for degree questions based on intervals: Negative islands and their obviation. *Journal of Semantics* 28:107–147.
- Szabolcsi, Anna. 1997b. Quantifiers in pair-list readings. In *Ways of scope taking*, ed. Anna Szabolcsi, 311–342. Dordrecht: Kluwer.

- von Stechow, Arnim, and Thomas Ede Zimmermann. 1984. Term answers and contextual change. *Linguistics* 22:3–40.
- Theiler, Nadine, Floris Roelofsen, and Maria Aloni. 2018. A uniform semantics for declarative and interrogative complements. *Journal of Semantics* 35:409–466.
- Uegaki, Wataru. 2015. Interpreting questions under attitudes. Ph.D. Dissertation, MIT.
- Uegaki, Wataru. 2018. On the projection of the presupposition of embedded questions. In *Proceedings of Semantics and Linguistic Theory 28*, ed., Sireemas Maspong, Brynhildur Stefánsdóttir, Katherine Blake, and Forrest Davis, 789–808. Washington, D.C.: LSA. <https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/28.789>.
- Uegaki, Wataru. 2021. The existential/uniqueness presupposition of *wh*-complements projects from the answers. *Linguistics and Philosophy* 44:911–951.
- Wagner, Michael, Elise McClay, and Lauren Mak. 2013. Incomplete answers and the rise-fall-rise contour. In *Proceedings of the Seventeenth Workshop on the Semantics and Pragmatics of Dialogue (SemDial)*, ed., R. Fernández and A. Isard, 140–149. http://events.illc.uva.nl/semDial/dialdam/papers/WagnerEtal_dialdam.pdf.
- Xiang, Yimei. 2016. Interpreting questions with non-exhaustive answers. Ph.D. Dissertation, Harvard University.
- Xiang, Yimei, and Alexandre Cremers. 2017. Mentions-some readings of plural-marked questions: Experimental evidence. In *Proceedings of North East Linguistics Society 47*, ed., Andrew Lamont and Katerina A. Tetzloff, volume 3, 261–274.
- Xiang, Yimei. 2019. Getting quantifying-into questions uniformly: Functionality, domain exhaustivity, and quantificational variability. In *Proceedings of Semantics and Linguistic Theory 29*, ed., Katherine Blake, Forrest Davis, Kaelyn Lamp, and Joseph Rhyne, 160–179. Washington, D.C.: LSA. <https://doi.org/10.3765/salt.v29i0.4610>.
- Xiang, Yimei. 2020. Function alternations of the Mandarin particle *dou*: Distributor, free choice licenser, and ‘even’. *Journal of Semantics* 37:171–217.
- Xiang, Yimei. 2021a. A hybrid categorial approach to question composition. *Linguistics and Philosophy* 44:587–647.
- Xiang, Yimei. 2021b. Higher-order readings of *wh*-questions. *Natural Language Semantics* 29:1–45.
- Xiang, Yimei. 2021c. Quantifying into *wh*-dependencies: Multiple-*wh* questions and questions with quantifiers. <https://ling.auf.net/lingbuzz/005097>, manuscript, Rutgers University.
- Zhao, Zhuoye. 2019. Varieties of distributivity: From Mandarin *dou* to plurality, free choice and scalarity. Master thesis. Institute for Logic, Language and Computation (ILLC), University van Amsterdam.