

Relativized Exhaustivity: Mention-Some and Uniqueness

Yimei Xiang, Rutgers University

(April 2, 2022)

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

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 attendees 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 ‘.../’).¹ Answers like (2a) are called *partial 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%

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 lack of final fall in (i) (marked by ‘...’) 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 ...’

- b. # 'Antonio did.'
L H* L-L% \rightsquigarrow *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 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 2016). 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 affects 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. 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*.

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 (EP) from Dayal 1996: a question is defined only if it has a unique exhaustive true answer. This presupposition, henceforth called 'Dayal's EP', nicely explains the uniqueness effects of singular *which*-phrases. To solve this dilemma, I will propose to replace Dayal's EP with a presupposition of '*Relativized Exhaustivity*', which has the effect of evaluating exhaustivity relative to the accessible worlds as opposed to the 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. Section 3 reviews existing approaches to the MS phenomenon and discusses the 'mention-one-only' constraint. 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 existing analyses. Section 6 proposes the concept of Relativized Exhaustivity, shows how it solves the dilemma, and accounts for local uniqueness. Section 7 concludes.

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 the utterance world. In the Kratzerian theory of modality, modal verbs vary along two axes, namely, *modal flavor* and *modal force*. 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 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 relevant circumstances, there is a world where we satisfy our goals by 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 2016: 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). The exhaustive answer is preferred because it maximizes the chance that the place that John actually left for is among the options.

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

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

A: a. ? 'Starbucks.'

b. 'Starbucks or Peet's.'

MS interpretations are also difficult in questions with a deontic modal. In (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 in order to maximize the information of relevance to Alice’s situation, the exhaustive answer (6b) is needed.

- (6) (Alice needs 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)

Modal force concerns the force of quantification. Only existential modals may license MS. This requirement will be a focus of this paper. In the examples below, the modal verb *should* is goal-oriented but has a universal modal force. In (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. If the quantification domain of *should* contains only 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. In English, these modalities are realized as *can* or through the use of infinitives (as in *where to get coffee*). In this paper, ‘*can*-questions’ refers to *wh*-questions with an existential teleological/bouletic modal.

2.2. Conversational goals

Conversational goals play an important role in distributing MS interpretations for matrix questions. On the one hand, questions without teleological/bouletic *can* admit MS answers in — and only in — goal-oriented contexts (Dayal 2016). (9) and (10) exemplify the case with the epistemic modal *might*: with an explicit conversational goal in the context, the answer in (10) can be read non-exhaustively. 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.*

- (10) a. I need help. Who might be in the office at this time?

- b. Bill might be in. \ (He could help you.)
 ↗ *The available evidence only suggests that BILL might be in.*
- (11) a. I need a ride to the party tonight. Who’s driving?
 b. Bill is. \ (He could give you a ride.)
 ↗ *Only Bill is driving.* ((9)–(11) are modified from Dayal 2016: 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) (The hiring committee wants to prioritize the candidates on the long-list who can teach Psycholinguistics.)
 Q: ‘Who can teach Psycholinguistics?’
 A: ‘Judy can. \’ \rightsquigarrow *Among the candidates on the long-list, only Judy can teach Psycholinguistics.*

The MS-licensing effect of goal-oriented contexts makes it appealing to think of MS as a pragmatics-driven phenomenon: MS is primarily licensed by a non-exhaustive conversational goal, independent of the presence of a modal expression. We could dub this the ‘modal-to-context reduction’ strategy: the reason why *can* licenses MS is that it is capable of conveying the goal-oriented modality.

In contrast, taking *can* as the primary MS-licensing factor, I analyze the MS-licensing effect of goal-oriented contexts in terms of ‘context-to-modal reduction’: these contexts may provide a constraint on the question nucleus, 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 x is in the office?
 b. Which x is such that x is driving and that x can give me a ride if x is driving?

With this strategy, whatever explains the MS-licensing effect of *can* also explains the MS-licensing effect of goal-oriented contexts.

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. 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.

Notably, regarding 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, they predict that any interrogative is ambiguous between MS and MA. In contrast, nucleus-dependent approaches draw close attention to the connection between the availability of MS interpretations and the presence of *can*.

3.1. Pragmatic approaches

MS was initially perceived as a pragmatic phenomenon. Earlier works adopted this view to maintain the core assumption that question interpretations are exhaustive (Groenendijk and Stokhof 1984).² MS answers, which are non-exhaustive, were therefore treated as partial answers permitted due to pragmatic considerations, such as their being sufficient relative to the conversational goals. Ginzburg (1995) and van Rooij (2003) assume that the semantics of an interrogative is underspecified, and that whether an answer is ‘complete’ is determined by how well this answer resolves the question relative to the goals of the questioner. van Rooij (2004) further relates question-answering to *decision problems* and develops a *utility* theory of answers.

The most commonly raised challenge to pragmatic approaches, as pointed out by Groenendijk and Stokhof themselves and reiterated by George (2011), has been that MS interpretations are also available in embeddings.³ As seen in (14), knowing a *can*-question implies knowing a MS answer to this question.

- (14) John knows [who can address this question].
 \rightsquigarrow For an individual *x* such that *x* can address this question, John 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. Thus, 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 (14). 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. 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 why *can* (or an equivalent modal expression) appears to be the only MS-licenser in embeddings. Dayal (2016: Chap. 3) observes that MS interpretations are less readily available in embeddings of a non-*can*-question. This observation has been experimentally validated: all else being equal, the absence of *can* in questions embedded under *remember* significantly decreases the acceptance of a MS interpretation (Xiang and Cremers 2017), and similarly for embeddings of non-finite questions (Moyer and Syrett 2019). Relatedly, non-interrogative *wh*-constructions such as *wh*-free relatives also require the presence of an existential goal-oriented modal to license an existential interpretation (Chierchia and Caponigro 2013). According to the aforementioned ‘context-to-modal reduction’ strategy, one plausible explanation for the contrast between matrix questions and embeddings in admitting MS interpretations is that

²To be exact, Groenendijk and Stokhof (1984) assume that question interpretations are *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* denotes a function that maps a world *w* to the exhaustified proposition ‘only *x* came’ that is true in *w*. However, most recent works on question semantics take the *weakly exhaustive* meaning, which only affirms the true answers, as the basic meaning of an interrogative. For example, *who came* primarily denotes a function that maps a world *w* to the proposition ‘*x* came’ such that it’s true in *w* that only *x* came. Strong exhaustivity, then, is derived from weak exhaustivity via a separate operation.

³Not all question embeddings admit MS readings. For example, *depends on*-sentences (e.g., *Where you can buy coffee depends on the time*) only allow for exhaustive readings (Groenendijk and Stokhof 1984; Dayal 2016).

the contextual support for licensing MS or other non-exhaustive interpretations is less accessible in embeddings.

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

I call an approach to MS ‘semantic’ if it (i) perceives MS as an independent interpretation of questions and (ii) attributes the MS/MA ambiguity to operations in the semantic composition of an interrogative. This classification isn’t rigid. For example, an unpublished version of Dayal 2016 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 ‘pragmatic’-vs-‘semantic’ demarcation 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. This brings us to the demarcation between nucleus-independent and 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 nature of the 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 two of these approaches.

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 in this world) but assume the availability of multiple answerhood operators, including:

$$(15) \quad \begin{aligned} \text{a. } \llbracket \text{ANS}_{BR1} \rrbracket &= \lambda w \lambda Q_{\langle s, \langle st, t \rangle \rangle} \cdot \bigcap \{ p \mid Q(w)(p) \wedge p(w) \} \\ \text{b. } \llbracket \text{ANS}_{BR3} \rrbracket &= \lambda w \lambda Q_{\langle s, \langle st, t \rangle \rangle} \lambda P_{\langle s, stt \rangle} \cdot \exists p [P(w)(p) \wedge Q(w)(p) \wedge p(w)] \end{aligned}$$

The above two answerhood operators differ in exhaustivity. In (16a), 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 interpretation. In contrast, in (16b), applying ANS_{BR3} to the embedded question returns an existential quantifier over questions. Interpreting this quantifier over the embedding predicate yields a MS interpretation.

$$(16) \quad \begin{aligned} \text{a. MA: } & [x \text{ knows } [\text{ANS}_{BR1}(w)(Q)]] \\ \text{b. MS: } & [\text{ANS}_{BR3}(w)(Q) \lambda p_{\langle s, t \rangle} \lambda w' [x \text{ knows}_{w'} p]] \end{aligned}$$

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 (17), 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 \times optionally acts on *Abs*. When \times is absent, the root denotes a set of non-exhaustified propositions as in (17d), each of which is a non-strongly-exhaustive (viz., MS or weakly exhaustive)

answer. When \mathbb{X} is present, the root denotes a set of exhausted propositions read as ‘Only the members of β came’ as in (17e), each of which is a strongly exhaustive answer.

(17) Who came?

a. $[[\text{Abs}]] = \lambda w \lambda x_e . \text{came}_w(x)$

b. $[[\text{Q}]] = \lambda \alpha_{\langle s, \tau t \rangle} \lambda p_{\langle s, t \rangle} . \exists \beta_\tau [p = \lambda w . \alpha(w)(\beta)]$

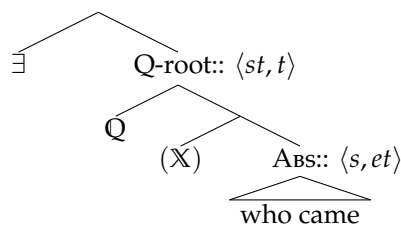
c. $[[\mathbb{X}]] = \lambda \gamma_\tau \lambda \delta_\tau [\delta = \gamma]$

d. Without \mathbb{X} : MS/ weakly exhaustive

$$[[\text{Q}(\text{Abs})]] = \lambda p_{\langle s, t \rangle} . \exists \beta_e [p = \lambda w . \text{came}_w(\beta)] \\ = \{ \lambda w . \text{came}_w(\beta) \mid \beta \in D_e \}$$

e. With \mathbb{X} : strongly exhaustive

$$[[\text{Q}(\mathbb{X}(\text{Abs}))]] = \lambda p_{\langle s, t \rangle} . \exists \beta_{\langle e, t \rangle} [p = \lambda w [(\lambda x_e . \text{came}_w(x)) = \beta]] \\ = \{ \lambda w . [(\lambda x_e . \text{came}_w(x)) = \beta] \mid \beta \in D_{\langle e, t \rangle} \}$$



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 \mathbb{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. Since the modal verb *can* appears inside the nucleus, only nucleus-dependent approaches may predict a grammatical relation between the presence of *can* and the availability of MS. 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.

MS as a scopal effect of exhaustification Distinct from the analysis reviewed in (17), George (2011: Chap. 6) assumes that the \mathbb{X} -operator is mandatorily used in question formation and treats MS as a scopal effect of this operator. When the \mathbb{X} -operator scopes below an existential expression, such as the existential modal *can*, the root denotes a set of propositions that are globally non-exhaustive.

(18) Who can chair the committee?

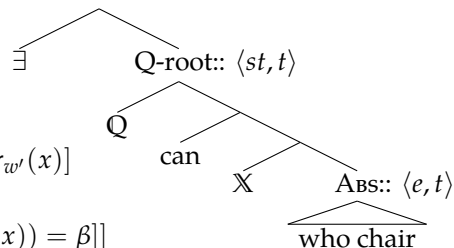
a. $[[\text{Abs}]] = \lambda x_e . \text{chair}_{w'}(x)$

b. $[[\mathbb{X}(\text{Abs})]] = \lambda \delta_{\langle e, t \rangle} [\delta = \lambda x_e . \text{chair}_{w'}(x)]$

c. $[[\text{can}]] = \lambda q_{\langle s, t \rangle} . \exists w' \in M_w [q(w')]$

d. $[[\text{can}(\mathbb{X}(\text{Abs}))]] = \lambda \delta_{\langle e, t \rangle} . \exists w' \in M_w [\delta = \lambda x_e . \text{chair}_{w'}(x)]$

e. $[[\text{Q}(\text{can}(\mathbb{X}(\text{Abs})))]] \\ = \lambda p_{\langle s, t \rangle} . \exists \beta_{\langle e, t \rangle} [p = \lambda w . \exists w' \in M_w [(\lambda x_e . \text{chair}_{w'}(x)) = \beta]] \\ = \{ \lambda w . \exists w' \in M_w [(\lambda x_e . \text{chair}_{w'}(x)) = \beta] \mid \beta \in D_{\langle e, t \rangle} \}$



This account predicts that MS interpretations are only available in questions with an existential expression. Moreover, with the local \mathbb{X} -operator, it also captures the ‘local-exhaustivity effect’ of MS answers. Compare the answers in (19): although it’s true that Andy can serve on the committee, (19a) is a bad answer, in contrast to (19b,c), each of which specifies the full composition of a possible

committee. This contrast argues that MS answers involve exhaustivity under *can*; for example, (19b) is read as: ‘It can be the case that *only* Andy and Billy serve on the committee.’

- (19) (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.\

This account also faces several problems. First, in composition, *can* selects for a proposition, but $\mathbb{X}(\text{Abs})$ denotes a set of predicates. Second, the global application of \mathbb{X} forces strong exhaustivity, which is too strong (see fn. 2). Third, this account predicts that any existential expressions, including indefinites, can license MS. However, a number of distinctions between *can*-questions and questions with an indefinite argue that their non-exhaustive interpretations have different origins (see fn. 18).

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 (20), a true answer is complete as long as it is not asymmetrically entailed by any other true answers. Here Q refers to the answer space (viz., the Hamblin set) of a question.

$$(20) \text{ANS}_{\text{Fox}} := \lambda w \lambda Q. \{p \mid w \in p \in Q \wedge \forall q [w \in q \in Q \rightarrow q \not\subseteq p]\} \quad (\text{Fox 2013})$$

This concept of answerhood is stronger than simple existentiality but weaker than exhaustivity. For example, it allows (18)/(21) to have multiple complete true answers of the form ‘*x* can chair’.

Fox (2013) further analyzes the MS/MA ambiguity as a scopal effect of distributivity. He assumes that the *wh*-trace and a covert distributivity operator form a distributive phrase [$X \text{ EACH}$]. If this phrase scopes below *can*, as in (21a), the answer space of the question is not closed under conjunction, so that applying ANS_{Fox} may return a set consisting of multiple MS answers. In contrast, if distributivity scopes above *can*, 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.

- (21) Who can chair the committee?
- a. $[_{\text{CP}} \text{ who } \lambda X \dots [_{\text{IP}} \text{ can } [[X \text{ EACH}] \lambda x [x \text{ chair the committee }]]]]]]$ (*can* \gg *EACH*: MS)
 - b. $[_{\text{CP}} \text{ who } \lambda X \dots [_{\text{IP}} [X \text{ EACH}] \lambda x [\text{ can } [x \text{ chair the committee }]]]]]]$ (*EACH* \gg *can*: MA)

Compared with George’s account, Fox’s account allows for weak exhaustivity. However, it doesn’t explain the local-exhaustivity effect and under-generates MS answers: for (19), it incorrectly predicts that (19b) is as bad as (19a), since they both are asymmetrically entailed by (19c). In addition, just like George’s, Fox’s account over-predicts a MS-licensing effect for indefinites. This over-generation problem also applies to the recent account in 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. They predict that the modal verb *can* (or an equivalent modal expression), which appears within the question nucleus, is the primary source for licensing MS. These approaches haven’t said much about the role of conversational goals in licensing MS, but they are compatible with the ‘context-to-modal reduction’ strategy laid out in Sect. 2.2. Other pragmatic factors may be treated as blockers for MS answers or rescuers for incomplete answers.

3.3. The ‘mention-one-only’ constraint

Pragmatic approaches and nucleus-independent approaches predict that the modal verb *can* may license MS just because it naturally comes with an existential conversational goal. A problem of this prediction concerning embeddings of non-*can*-questions has been discussed in Sect. 3.1. As for why conversational goals may license MS, pragmatic approaches attribute this 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, the non-exhaustive interpretations of *can*-questions and non-*can*-questions have the same nature. In what follows, I will argue against this prediction based on a new observation on MS answers, which cannot be explained by pragmatics.

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. Notably, *mention-few* differs from *mention-one* in two aspects: (i) embeddings of *can*-questions admit *mention-one* but not *mention-few* interpretations; (ii) in discourse, unlike *mention-one* answers, *mention-few* answers to *can*-questions easily imply exclusivity if not ignorance-marked.

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*. Consider (22), which involves a polar question with embedded *can*. 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 “Could you tell me three places on campus where one can hold a large conference?”⁴

- (22) 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 large conference?’
Bob: ‘Yes, but I only know one such place.’/‘#No/Sorry, I only know one such place.’

Second, unlike *mention-one* answers, *mention-few* answers are not read non-exhaustively in the absence of an ignorance marker. In an informal comprehension task I posted on social network sites in both English and Mandarin, speakers were asked to judge whether either the *mention-one* answer (23a) or the *mention-two* answer (23b) gave rise to an exclusivity inference. Notably, among those who judged (23a) non-exhaustively, a large majority judged (23b) exhaustively. The contrast in exhaustivity between (23a) and (23b) suggests that there is a crosslinguistic grammatical constraint that only allows *mention-one* answers to be MS answers.⁵

- (23) (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.)

⁴Note that in this scenario it is infelicitous to ask “Do you know *which three places* on campus one can hold a large conference?”, because this question presupposes that there are only three such places on campus.

⁵One might give a pragmatic explanation as follows: (23b) is read exhaustively because it is over-informative, in contrast to (23a), which is just as informative as needed. However, it remains puzzling why over-informativity forces exhaustivity. One would need to assume two distinct answerhood operations: (i) a pragmatic answerhood that seeks a (possibly non-exhaustive) answer that is as informative as needed, and (ii) a semantic answerhood that demands an exhaustive answer. Nevertheless, this hybrid analysis cannot explain (22): when the questioner’s goal is ‘*mention-three*’, a *mention-one* answer wouldn’t be permitted by any answerhood operations — it is neither informative enough nor exhaustive.

- Bob: ‘Which blade can I use to cut carrots?’
- Alice: a. ‘The green one.’ (Mention-one: likely non-exhaustive)
 b. ‘The green one or the black one.’ (Mention-few: likely exhaustive)

For answers to questions with a partiality marker (e.g., *for example*), there is no contrast in exhaustivity between mention-one and mention-few: in (24), neither answer implies exclusivity. What’s more, the questioner can make an exact ‘mention-*N*’ inquiry by explicitly saying “Give me *N* examples”.

- (24) Who is on your committee, for example?
- a. Andy is on my committee. \ (Mention-one: non-exhaustive)
 b. Andy and Billy are on my committee. \ (Mention-few: non-exhaustive)

The contrast between the *can*-question (23) and the *for example*-question (24) 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 (Xiang 2021a). In contrast, the non-exhaustivity of *can*-questions is grammatically obtained from the answerhood operator for complete true answers.

Finally, let me clarify a possible point of confusion: ‘mention-one’ refers to mentioning one option, not one atomic individual. An option can be made up of either an atomic individual or the sum or a Boolean coordination of multiple individuals. For example, a mention-one answer to (19) *Who can serve on the committee?* should specify the sum of a group of individuals who can simultaneously serve on the committee. Similarly, the answer in (25), which names a Boolean conjunction, expresses a single option to assign leaders.

- (25) (The players can be grouped into 2–3 teams. Each team needs one or two leaders.)
- Q: ‘Who can we ask to lead a team?’
 A: ‘We can ask Alex to lead a team and Ben to lead another team.’

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, 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/higher-order MS (Sect. 4.2), conjunctive MA (Sect. 4.3), and disjunctive MA (Sect. 4.4).

4.1. General assumptions on questions and answers

4.1.1. Questions as topical properties

I define the root denotations of questions as *topical properties* (Chierchia and Caponigro 2013).⁶ As exemplified in (26a,b), the topical property of a *wh*-question is a function that maps an individual in the *wh*-domain (viz., a *short answer*) to a proposition in the answer space (viz., a *propositional answer*). In other words, as formalized in (26c), the answer space of a question is the *image* (i.e., the set of all

⁶This paper stays neutral on how topical properties are compositionally derived. I will only specify the composition of the question nucleus. Besides categorial approaches, which define questions as functions, dynamic approaches (Dotlačil and Roelofsen 2019; Li 2021) also allow for extraction of short answers from question denotations.

output values) of the topical property of this question. I henceforth denote a topical property as ‘ $\llbracket Q \rrbracket$ ’ and an answer space as ‘ Q ’.

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

Defining questions as topical properties, rather than as partitions of worlds or sets of propositions, makes it easy to track short answers, whose semantics are modal independent. This analytical choice will be crucial for modeling ‘Relativized Exhaustivity’ (see fn. 21), a question interpretation condition that I propose to solve the dilemma between uniqueness and MS.

4.1.2. Answerhood

Many theories developed in the past three decades (Heim 1994; Dayal 1996; a.o.) encode the question properties of truth and 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 exhaustivity.

Adopting the concept of answerhood in Fox 2013, I assume that there is only one type of exhaustivity grammatically encoded in answerhood operators, henceforth called ‘*maximal (max-)informativity*’. Max-informativity is stronger than simple existentiality but weaker than exhaustivity. A true answer is called ‘max-informative’ if and only if it is not asymmetrically entailed by any other true answers.

Adapting the ANS_{Fox} -operator in (20) to the assumed question semantics, I define a pair of answerhood operators as in (27). These operators are applied to a topical property, not a Hamblin set. The superscripts S and P stand for ‘short’ and ‘propositional’, respectively. $\llbracket Q \rrbracket_w := \{ \alpha \mid \alpha \in \text{Dom}(\llbracket Q \rrbracket) \wedge w \in \llbracket Q \rrbracket(\alpha) \}$ abbreviates the set of short answers to Q that are true in w .

- (27) Answerhood operators for complete true answers (to be modified)
- a. $\text{ANS}^S(w)(\llbracket Q \rrbracket) = \{ \alpha \mid \alpha \in \llbracket Q \rrbracket_w \wedge \forall \beta [\beta \in \llbracket Q \rrbracket_w \rightarrow \llbracket Q \rrbracket(\beta) \not\subset \llbracket Q \rrbracket(\alpha)] \}$
 - b. $\text{ANS}^P(w)(\llbracket Q \rrbracket) = \{ \llbracket Q \rrbracket(\alpha) \mid \alpha \in \text{ANS}^S(w)(\llbracket Q \rrbracket) \}$

4.1.3. First-order versus higher-order interpretations

Wh-questions are ambiguous between first-order and higher-order interpretations. For example, the *have to*-question in (28) can be completely addressed by specifying particular books, as in (28a), or by using a generalized quantifier (GQ) over a set of books, as in (28b) (Spector 2007).

- (28) 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 LFs and denotations for the question in (28) as follows:

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

In LF (30a), the *wh*-phrase binds a higher-order trace π across the modal verb (after Spector 2007, 2008). The interactions between this 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 a covert scopal expression inside the question nucleus — in contrast to George 2011 and Fox 2013, which analyze MS as a scopal effect of exhaustivity or distributivity. In denotation (30b), ‘ ${}^H \mathbf{books}_@$ ’ 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 (Spector 2007, 2008; Xiang 2021b), for this paper all that matters is that the domain of a higher-order *wh*-quantification 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 an LF like (31). 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).

- (31) Who can serve on the committee? (First-order MS)
 $[_{CP} \text{ who } \lambda x_e [_{IP} \text{ can } [O_C [_{VP} x \text{ serve on the committee }]]]]$

The local O -operator is assumed to capture the local exhaustivity and mutual independence effects of MS answers. In (32), repeated from (19), local exhaustivity says that a MS answer should specify all the members of a possible committee. Mutual independence is a property predicted by Fox’s (2013) definition of answerhood: (32b) counts a good MS answer only if it isn’t asymmetrically entailed by (32c).

- (32) (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.) (= (19))

Who can serve on the committee?

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

As defined in (33), the O -operator affirms the prejacent proposition and negates the alternatives of the prejacent that are not entailed by that prejacent. The domain variable C carried by the O -operator denotes a contextually determined subset of the alternatives.

- (33) $\llbracket O_C \rrbracket = \lambda p \lambda w . p(w) = 1 \wedge \forall q \in C [p \not\subseteq q \rightarrow q(w) = 0]$ (Chierchia et al. 2012; a.o.)

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 (17)); however, as demonstrated below, the O -operator is technically neater in composition.

The LF (31) is now computed as follows. Here the O -operator is associated with an e -type wh -trace x . This trace is associated with a set of ‘variable alternatives’, defined like focus alternatives as in (35).⁷ The domain variable C carried by the O -operator denotes a contextually determined subset of the variable alternatives of the VP. Composing this LF yields the topical property (34b). In the context described in (32), applying the assumed answerhood operator returns a set with two members, given in (34c), each of which is a MS answer.

- (34) $[\text{CP who } \lambda x_e [\text{IP can } [O_C [\text{VP } x \text{ serve on the committee }]]]]$ (= (31))
- a. $v\text{-ALT}(\text{VP}) = \{\phi_x \mid x \in D_e\}$ (ϕ_x abbreviates ‘ x serves on the committee’)
- b. $[\text{Q}] = \lambda x_e : x \in \text{hmn}_@ \cdot \diamond O_C \phi_x$, where $C \subseteq \{\phi_x \mid x \in \text{hmn}_@\}$
- c. $\text{ANS}^P(w)([\text{Q}]) = \{\diamond O_C \phi_{a \oplus b}, \diamond O_C \phi_{a \oplus b \oplus c}\}$
- (35) For any trace and pronoun α , we have: $v\text{-ALT}(\alpha) = D_{\text{type}([\alpha])}$.

I assume that the insertion of the local O -operator is optional, given that the local exhaustivity implicature is cancellable and suspendable, as seen in (36). This observation also argues that mandatory effects, such as the uniqueness effects of singular wh -questions and the unavailability of MS in non-*can*-questions, are independent of the presence of this O -operator. Section 6 will account for these mandatory effects without resorting to a local O -operator.

- (36) Who can serve on the committee?
- a. Andy and Billy ... maybe also Cindy.
- b. 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 (37). To avoid complications from plural answers, the example sentence uses the predicate *chair the committee*. Compared with the derivation of the first-order MS interpretation in (34), 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 between *can* and the local O -operator. Composing this LF yields the topical property (37b). (^Hhmn_@ stands for a set of GQs over human individuals.)

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

Figure 1 illustrates the answer space of (37). This illustration contains four answers related to the two individuals Andy (a) and Billy (b), derived by applying the topical property (37b) to the Montagovian individuals a^\uparrow and b^\uparrow , their Boolean conjunction $a^\uparrow \cap b^\uparrow$, and their Boolean disjunction $a^\uparrow \cup b^\uparrow$.⁸ 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 ’.

⁷Unlike focus alternatives, variable alternatives are specific to variable-like expressions and do not require focus marking.
⁸For any meaning α of type τ , we have: the Montague-lifted meaning of α is α^\uparrow (of type $\langle \tau t, t \rangle$) s.t. $\alpha^\uparrow := \lambda m_{\langle \tau, t \rangle} \cdot m(\alpha)$. 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.)

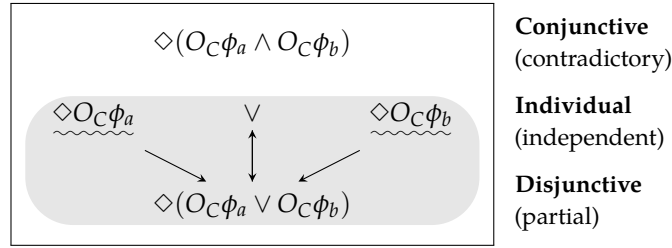


Figure 1: The answer space of the higher-order MS interpretation of *Who can chair the committee?* [Arrows indicate entailment relations, shading marks the true answers, and underwaving marks the max-informative true answers. The bi-implication between ‘ \vee ’ in the middle and the disjunctive answer 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 are characterized as follows:

- *Conjunctive answers*: Due to the exclusivity of the local O -operator, all conjunctive answers are contradictory. For example, $\diamond(O_C\phi_a \wedge O_C\phi_b)$ means ‘#There is a world w compatible with the current circumstances such that only a chairs in w and only b chairs 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 like $\diamond O_C\phi_{a\oplus b}$, no matter whether these answers are read as distributive or non-distributive. Hence, any true individual answer is a max-informative true answer.
- *Disjunctive answers*: 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 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, which each specify one 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 coordination to specify multiple options, she must be understanding the question with a MA interpretation, which calls for an exhaustive answer.

4.2.3. Predictions on the distribution of MS

According to Fox’s definition of answerhood, a question has a MS interpretation 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 *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 (38b) or replaced with a universal modal as in (38c), the individual answers are mutually exclusive.

(38) Let $C = \{\phi_x \mid x \in D\}$. For any a and b in D s.t. $\phi_a \neq \phi_b$, we have:

- a. $\diamond O_C \phi_a \wedge \diamond O_C \phi_b \not\equiv \perp$
- b. $O_C \phi_a \wedge O_C \phi_b \equiv \perp$
- c. $\Box O_C \phi_a \wedge \Box O_C \phi_b \equiv \perp$

For condition (B), as seen in Figure 1 above, the answer space of a *can*-question with a higher-order interpretation is not closed under conjunction if the higher-order *wh*-trace scopes below *can*. In contrast, for a *wh*-question without a modal or with a universal modal, the answer space derived in a higher-order interpretation is closed under conjunction: in each illustration in Figures 2a–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).

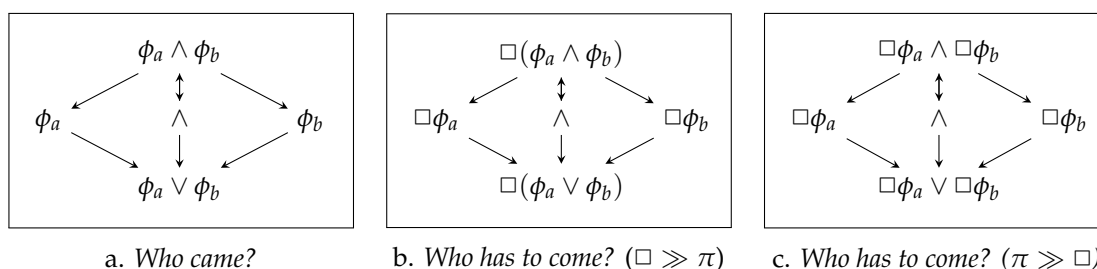


Figure 2: Answer spaces of non-modalized questions and *have to*-questions

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 instance, consider the non-modalized *wh*-question in (39), which has a non-distributive predicate, *form a team*. Although the answer space yielded in the higher-order (HO) interpretation is closed under conjunction, the answers allowed in the first-order (FO) interpretation are logically independent. Hence, unless we can independently rule out (39a), the assumptions made in this section would predict a MS/MA ambiguity for (39), contrary to fact.

- (39) (The children formed two teams: $a + b$ formed one, and $c + d$ formed the other.)
 Which children formed a team? (ϕ_x abbreviates ‘ x formed a team’)
- a. True answers to Q_{FO} : $\{\phi_{a \oplus b}, \phi_{c \oplus d}\}$
 - b. True answers to Q_{HO} : $\{\phi_{a \oplus b}, \phi_{c \oplus d}, \phi_{a \oplus b} \wedge \phi_{c \oplus d}, \phi_{a \oplus b} \vee \phi_{c \oplus d}\}$

There are two ways to solve the MS over-generation problem in (39). 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 (see also fn. 19). 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 $\{\{a, b\}, \{c, d\}\}$. In this proposal, if higher-order pluralities are available in the first-order *wh*-quantification, the answer space of (39a) is closed under conjunction.

The other way is to rule out (39a) by an independent constraint. The MS over-generation problem arises not only in *wh*-questions with a non-distributive predicate but also in other questions (Sect. 5.1.2). It's more appealing to resolve these cases uniformly.

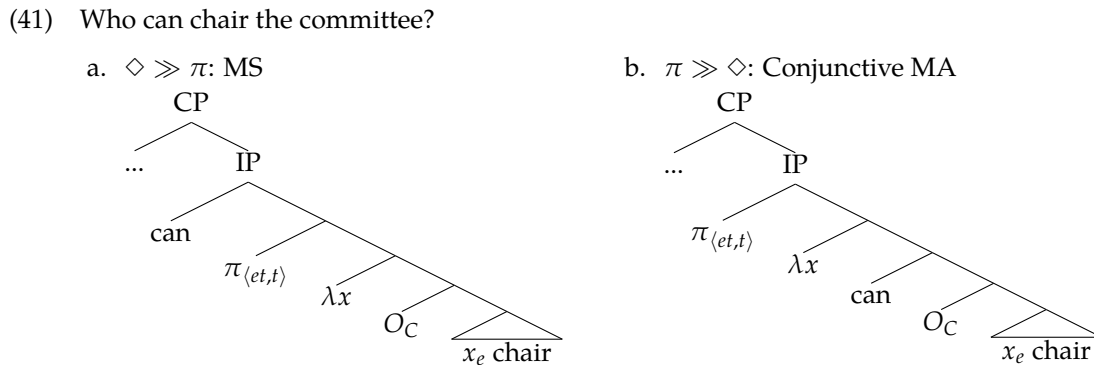
Note that conceptually the over-generation problem laid out in this section is the price we pay for abandoning Dayal's EP: if question interpretations had to be exhaustive, any interpretation that does not render the existence of an exhaustive true answer would be deviant. To predict the distribution of MS, I will argue below for a condition called 'Relativized Exhaustivity', which works like Dayal's EP 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 interpretations of the question, and that the two interpretations are compositionally derived from LFs with different question nuclei.

- (40) Who can chair the committee?
 a. Andy can, and Billy can. (Conjunctive MA)
 b. Andy or Billy can. (Disjunctive MA)

The derivation of the conjunctive MA interpretation is straightforward: it arises if the higher-order *wh*-trace scopes above the existential modal. Compared with (41a), which is simplified from (37), the only change in (41b) is that the *wh*-trace π scopes above *can*.



The answer spaces yielded by (41a,b) are illustrated in Figures 3a,b, respectively. (Figure 3a repeats Figure 1.) In Figure 3b, the answer space is closed under conjunction. The conjunctive answer, which is logically equivalent to the conjunction of the two individual answers, is the unique max-informative true answer to the question. Hence, the resulting interpretation is conjunctive MA.

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

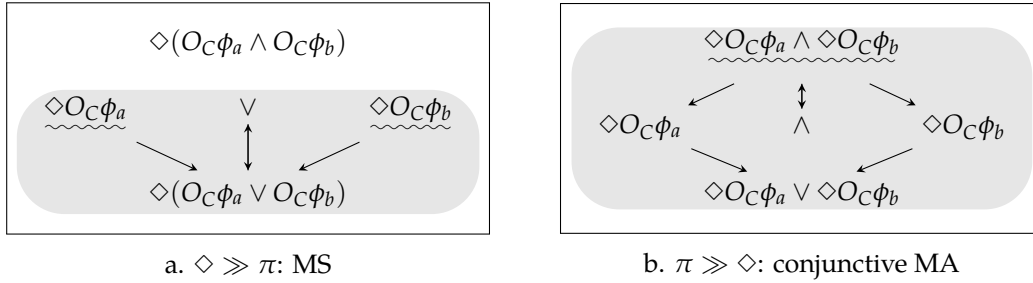


Figure 3: Answer spaces yielded by (41a,b) [Legends and abbreviations as in Figure 1.]

4.4. Deriving disjunctive MA interpretations

MA answers to *can*-questions are more commonly expressed as disjunctions rather than conjunctions. As exemplified in (42), 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.

- (42) Who can teach Calculus?
- a. Andy or Billy (can) ... (but I don't know who). (Ignorance: partial)
 \rightsquigarrow *Either Andy or Billy can teach Calculus, but I don't know which one.* ($\diamond\phi_a \vee \diamond\phi_b$)
- b. Andy or Billy (can). (FC: mention-few/all)
 \rightsquigarrow *Andy can teach Calculus, and Billy can teach Calculus, 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 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 is not always strongly exhaustive; and (ii) it explains why only *can*-questions admit FC-disjunctive answers. In the following, I will account for (i) by 'anti-exhaustification' and explain (ii) by 'Relativized Exclusivity'.

4.4.1. Deriving universal FC by anti-exhaustification

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 $\diamond\psi$ yields 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'.

The concept of anti-exhaustification has been widely implemented in exhaustification-based approaches to FC (Fox 2007; Chierchia 2013; Bar-Lev and Fox 2020; a.o.). For example, Fox (2007) analyzes anti-exhaustivity as a consequence of applying innocent-exclusion (IE-)based exhaustification recursively. This analysis predicts that FC and exclusivity arise at the same time. In contrast,

to allow disjunctive MA interpretations to be weakly exhaustive (fn. 2), I treat exhaustification and anti-exhaustification, which derive exclusivity and FC, respectively, as two independent operations.

Analogously to the exhaustification operator being called ‘*O*’, in reference to *only*, I here call the anti-exhaustification operator ‘*DOU*’, in reference to the Mandarin particle *dou*, which displays the properties of such operator overtly in questions and declaratives.⁹ *Dou* blocks MS in *can*-questions, as in (43), and triggers FC in *can*-disjunctions, as in (44). (Underlining marks the expression associated with *dou*.) Note that (44) doesn’t give rise to exclusivity: even with *dou* present, (44) doesn’t imply that no one else can teach Calculus.

- (43) (Dou) shui keyi jiao weijifen?
 (DOU) who can teach Calculus
 Without *dou*: ‘Who can teach Calculus?’ (✓MS, ✓MA)
 With *dou*: ‘Who all can teach Calculus?’ (✗MS, ✓MA)
- (44) Yuehan huozhe Mali (dou) keyi jiao weijifen.
 John or Mary (DOU) can teach Calculus
 Without *dou*: ‘Either John or Mary can teach Calculus.’ (Ignorance)
 With *dou*: ‘John and Mary (and possibly others) can teach Calculus.’ (Universal FC)

As schematized in (45), the *DOU*-operator I wish to postulate for English affirms the prejacent and negates the (IE-based) exhaustification of each ‘non-innocently-excludable alternative’ except the prejacent itself.

- (45) Anti-exhaustification operator *DOU*
 $\llbracket \text{DOU}_C \rrbracket = \lambda p \lambda w : p(w) = 1 \wedge \forall q \in (C - \{p\}) [q \notin \text{IExCL}(p, C) \rightarrow O_C^{\text{IE}}(q)(w) = 0]$

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* (Fox 2007). ‘IE-based exhaustification’ is the exhaustification operation that negates only the I-excludable alternatives (cf. the *O*-operator in (33), which negates all the non-entailed alternatives).

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

Innocent exclusion differs from traditional exclusion mainly in sentences with disjunctions or indefinites. A sentence with a disjunction is associated with a set of domain alternatives, derived compositionally from the domain alternatives of the disjunctive *or*:

- (47) $\text{D-ALT}(or) = \{\lambda b \lambda a . a \sqcup b, \lambda b \lambda a . a, \lambda b \lambda a . b\}$

For example, 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, $\diamond(\phi \vee \psi) \wedge \neg\diamond\phi \wedge$

⁹What is particularly interesting about treating *dou* as an anti-exhaustification operator is that it accounts for multiple uses of *dou*. As argued in Xiang 2020, the function of *dou* varies depending on what alternatives participate in anti-exhaustification. Assuming a non-vacuity presupposition on the existence of such alternatives, this analysis accounts for the universal distributor use, the FC-trigger use, and the *even*-like use. Alternatively, by giving *dou* an *even*-like semantics (Liu 2016), Mingming Liu derives the anti-exhaustivity effect indirectly (reviewed in Xiang 2020: Appendix B). For an account that derives the FC-trigger use of *dou* without assuming anti-exhaustification, see Zhao 2019.

$\neg \diamond \psi = \perp$). In contrast, for the narrow-scope \Box -disjunction $\Box(\phi \vee \psi)$, the subdomain alternatives $\Box\phi$ and $\Box\psi$ are I-excludable because $\Box(\phi \vee \psi) \wedge \neg\Box\phi \wedge \neg\Box\psi \neq \perp$.

I henceforth call the alternatives that participate in anti-exhaustification ‘anti-excludable alternatives’ (abbreviation: ANTIEXCL).¹⁰ The definition in (45) is re-written as follows:

$$(48) \quad \llbracket \text{DOU}_C \rrbracket = \lambda p \lambda w : p(w) = 1 \wedge \forall q \in \text{ANTIEXCL}(p, C) [O_C^{\text{IE}}(q)(w) = 0]$$

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

Now the alternatives are divided into two categories: I-excludable alternatives, which participate in exhaustification, and anti-excludable alternatives, which participate in anti-exhaustification. More examples are 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 [‘Neither’ means ‘neither weaker nor I-excludable’.]

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 definition (48), 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 (44) is now computed as in (49). Associated with a disjunction, DOU quantifies over the set of domain alternatives of its prejacent, as in (49b), which includes two anti-excludable alternatives, as in (49c). Employing DOU affirms the prejacent disjunctive sentence and negates the exhaustification of each disjunct, yielding a universal FC inference, as in (49d).

$$(49) \quad \llbracket \text{DOU}_C \llbracket_s \text{John or Mary can teach Calculus} \rrbracket \rrbracket$$

- a. $\llbracket S \rrbracket = \Diamond\phi_j \vee \Diamond\phi_m$ (ϕ_x abbreviates ‘ x teaches Calculus’)
- 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. $\llbracket \text{dou}_C(S) \rrbracket \Leftrightarrow [\Diamond\phi_j \vee \Diamond\phi_m] \wedge \neg O_C^{\text{IE}}\Diamond\phi_j \wedge \neg O_C^{\text{IE}}\Diamond\phi_m$
 $\Leftrightarrow [\Diamond\phi_j \vee \Diamond\phi_m] \wedge [\Diamond\phi_j \rightarrow \Diamond\phi_m] \wedge [\Diamond\phi_m \rightarrow \Diamond\phi_j]$
 $\Leftrightarrow [\Diamond\phi_j \vee \Diamond\phi_m] \wedge [\Diamond\phi_j \leftrightarrow \Diamond\phi_m]$
 $\Leftrightarrow \Diamond\phi_j \wedge \Diamond\phi_m$

4.4.2. 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,

¹⁰I 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 (see (72) in Sect. 5.2.1).

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 \Diamond -sentences (i.e., $\Diamond(\phi \vee \psi)$ and $\Diamond\phi \vee \Diamond\psi$).¹¹

In the realm of exhaustification-based theories of FC, explanations of modal obviation 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; a.o.): 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 give the wanted result in questions with a uniqueness effect. (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* presupposes *Relativized Exclusivity* (*RelExcl*) (cf. the *Viability* constraint of Dayal 2013).¹² This presupposition concerns the alternatives that participate non-vacuously in anti-exhaustification: for any two such alternatives $\Delta\phi$ and $\Delta\psi$ (Δ is null or a modal), if they are mutually excludable, their prejacent ϕ and ψ cannot be true in the same accessible world. In other words, $\Delta\phi$ and $\Delta\psi$ cannot be verified by the same set of accessible worlds.¹³

A formal definition of *RelExcl* is given in (50): for every alternative ϕ such that ϕ is anti-excludable and stronger than the prejacent, every minimal set of accessible worlds that verifies ϕ also verifies the (IE-based) exhaustification of ϕ . Here M is a modal base, and C/C' denotes a subset of the alternatives of $\llbracket S \rrbracket^M / \llbracket S \rrbracket^{M'}$:^{14,15}

$$(50) \text{ Relativized Exclusivity} \\ \llbracket \text{DOU}_C(S) \rrbracket^M \text{ is defined in } w \text{ only if}$$

¹¹The FC-trigger use of *dou* is also subject to modal obviation (Xiang 2020): 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.

- (i) a. Yuehan huozhe Mali dou keyi/*bixu jiao weijifen.
John or Mary dou can/*must teach Calculus
- b. Yuehan huozhe Mali (*dou) jiao -guo weijifen.
John or Mary (*dou) teach -EXP Calculus

¹²To be exact, although *RelExcl* is treated as a definedness condition of anti-exhaustification here, the effects of *RelExcl* are independent of whether universal FC is derived by 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.

¹³For a set of possible worlds W and a sentence ϕ , we have:

- (i) a. W verifies $\Diamond\phi$ in w iff there is a w -accessible world w' in W s.t. ϕ is true in w' ;
- b. W verifies $\Box\phi$ in w iff for every w -accessible world w' in W , ϕ is true in w' .

¹⁴In this definition, ϕ is a syntactic expression, not a semantic value. *RelExcl* is formalized as such because it concerns the interpretations of the alternatives evaluated relative to different modal bases.

¹⁵*RelExcl* 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 *Viability* analogously to (50):

- (i) $\forall\phi[\llbracket\phi\rrbracket^M \in (\text{ANTIEXCL}(\llbracket S \rrbracket^M, C) \wedge \llbracket\phi\rrbracket^M \subset \llbracket S \rrbracket^M \rightarrow \exists M'[M'_w \subseteq M_w \wedge O_C^E(\llbracket\phi\rrbracket^{M'})(w) = 1]]$
(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 (53a), 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.

$\forall \phi [\llbracket \phi \rrbracket^M \in \text{ANTIEXCL}(\llbracket S \rrbracket^M, C) \wedge \llbracket \phi \rrbracket^M \subset \llbracket S \rrbracket^M$
 $\rightarrow \forall M'_{\langle s, st \rangle} [M'_w \text{ is a minimal subset of } M_w \text{ s.t. } \llbracket \phi \rrbracket^{M'}(w) = 1 \rightarrow O_C^{\text{IE}}(\llbracket \phi \rrbracket^{M'})(w) = 1]$
 (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 RelExcl accounts for modal obviation. For simplicity, the formula $O_C^{\text{IE}}\phi$ (viz., the IE-based exhaustification of ϕ relative to C) will be abbreviated as $O\phi$.

For a non-modalized disjunction, the choice of modal base makes no difference. For (51), when the FC inference $\phi_j \wedge \phi_m$ is true, RelExcl requires both $O\phi_j$ and $O\phi_m$ to be true, causing a contradiction.

- (51) * DOUC [John or Mary teach Calculus]
 (Ungrammatical because $\text{DOUC}[\phi_j \vee \phi_m]$ inevitably violates RelExcl.)

For a wide-scope \Box -disjunction, there is no modal base that allows the universal FC inference and the RelExcl presupposition to be simultaneously true. For (52), we have: (i) the FC inference $\Box\phi_j \wedge \Box\phi_m$ is true in w relative to M if and only if John and Mary both teach Calculus in every world in M_w ; (ii) $O\Box\phi_j$ is true in w relative to M' if and only if only John teaches Calculus in every world in M'_w ; (iii) clearly M'_w cannot be a subset of M_w unless $M'_w = \emptyset$.

- (52) * DOUC [John or Mary must teach Calculus]
 (Ungrammatical because $\text{DOUC}[\Box\phi_j \vee \Box\phi_m]$ inevitably violates RelExcl.)

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 (53) 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 Calculus. As stated in (53), assume that only John, only Mary, and they both teach Calculus 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'' .

- (53) DOUC [John or Mary can teach Calculus]

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

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

If the sentence is parsed without local exhaustification, as in (53a), RelExcl is satisfied only in w . The minimal set of w -accessible worlds that verifies $\Diamond\phi_j$, namely $\{w_1\}$, also verifies the 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 (53b), RelExcl is also satisfied in w' : the violation of exclusivity in w_3 does not affect RelExcl because $\{w_3\}$ doesn't verify any of the

locally exhausted 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.

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

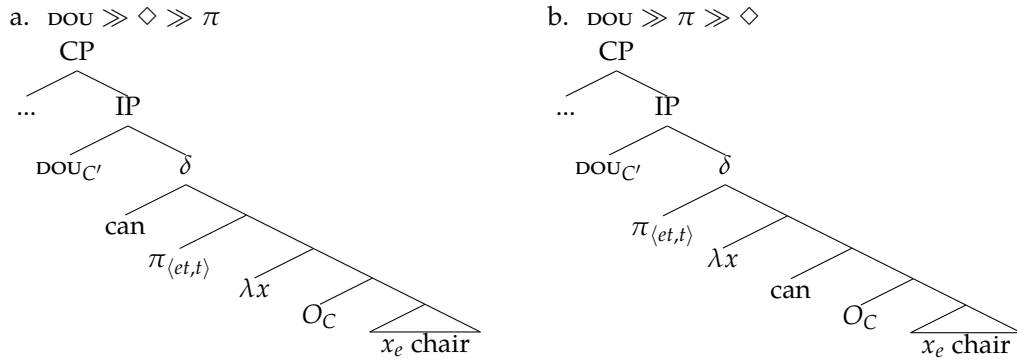
- (54) a. $\text{DOU}_C[\diamond(\phi_j \vee \phi_m)] = \diamond\phi_j \wedge \diamond\phi_m$, defined only if $\diamond O\phi_j \wedge \diamond O\phi_m \wedge \neg\diamond(\phi_j \wedge \phi_m)$.
 b. $\text{DOU}_C[\diamond(O\phi_j \vee O\phi_m)] = \diamond O\phi_j \wedge \diamond O\phi_m$, defined only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

To sum up, I derive universal FC by the anti-exhaustification operator DOU . This operator presupposes 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.

4.4.3. Deriving disjunctive MA

Prompted by the suggestive parallel uses of *dou* in disjunctions and *wh*-questions in Mandarin, I would like to 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. The resulting LFs are structured as in (55). The δ -nodes in (55a) and (55b) 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 *wh*-trace π relative to the modal verb *can*. In both LFs, a covert DOU -operator 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 .

- (55) ‘Who can chair the committee?’ (Disjunctive MA)
 ‘Andy or Billy can.’ \rightsquigarrow *Andy can chair the committee, and Billy can chair the committee.*



Compare now the answer spaces yielded in the absence of DOU , which are displayed once more in Figures 4a,b, to the answer spaces yielded by the LFs in (55a,b), depicted in Figures 5a,b. Legends, abbreviations, and the setup of the context are the same as in Figure 1.¹⁶

¹⁶Figure 5b is labeled as ‘dis/conjunctive 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.

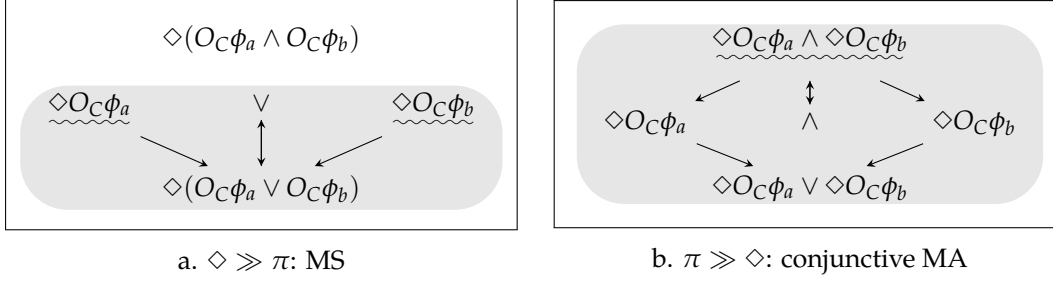


Figure 4: Answer spaces of (55a,b) yielded in the absence of DOU

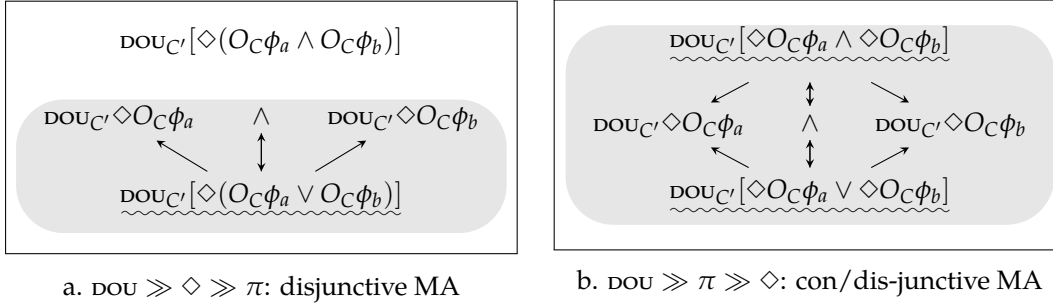


Figure 5: Answer spaces of (55a,b) yielded in the presence of DOU

The computation of the answers in Figure 5b is given in (56) below. 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 answer has anti-excludable alternatives (viz., the conjuncts), but these alternatives are weaker than the prejacent. However, as computed in (56c), 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. The computation of the alternatives in Figures 5a is similar, except that the conjunctive answer has a contradiction-denoting prejacent and thus is a contradiction itself.

- (56) For $DOU \gg \pi \gg \diamond$ (answer space in Figure 5b):
- a. $C = \{\phi_x \mid x \in \text{hmn}_\emptyset\}$, $C' = \{\pi(\lambda x_e. \diamond O_C \phi_x) \mid \pi \in {}^{\text{h}}\text{hmn}_\emptyset\}$
 - b. Conjunctive answer: wide-scope conjunction
 - 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. $DOU_{C'}[\diamond O_C \phi_a \wedge \diamond O_C \phi_b] \Leftrightarrow \diamond O_C \phi_a \wedge \diamond O_C \phi_b$
(Read as: 'A alone as well as B alone can chair the committee.')
 - c. Disjunctive answer: universal FC
 - 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. $DOU_{C'}[\diamond O_C \phi_a \vee \diamond O_C \phi_b] \Leftrightarrow \diamond O_C \phi_a \wedge \diamond O_C \phi_b$
(Read as: 'A alone or B alone DOU can chair the committee.' (or \gg can))

This analysis also 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 (28), repeated in (57), a disjunctive answer may express either ignorance ($\square \phi_f \vee \square \phi_r$)

or existential FC ($\Box[\phi_f \vee \phi_r]$), but not universal FC ($\Box\phi_f \wedge \Box\phi_r$): applying DOU to the disjunctive answer either inevitably violates RelExcl or is vacuous.

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

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 scopes above *can*, or (ii) an anti-exhaustification operator 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.

The proposed account captures several linguistic properties of MS answers, namely local exhaustivity, mutual independence, and ‘mention-one-only’. Second, by assuming higher-order *wh*-quantification, it derives a scopal effect in *can*-questions for free; this scopal effect naturally derives the contrast between MS and conjunctive MA. Third, by assuming anti-exhaustification and RelExcl , this account uniformly explains the derivation and distribution of universal FC disjunctions and disjunctive MA interpretations.

However, as noted 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 \Box -questions. As I will argue next, 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. However, this definition of answerhood conflicts with Dayal’s (1996) exhaustivity presupposition (EP), 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. 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 in *which*-questions

Wh-questions with a singular *which*-phrase (called ‘singular *wh*-questions’) are subject to a uniqueness effect. For example, question (58) is felicitous only if the speaker believes that this question has just one true answer.

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

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

$$(59) \text{ Ans}_{\text{Dayal}} := \lambda w \lambda Q : \exists p [w \in p \in Q \wedge \forall q [w \in q \in Q \rightarrow p \subseteq q]].$$

$$\text{ip} [w \in p \in Q \wedge \forall q [w \in q \in Q \rightarrow p \subseteq q]]$$

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*-quantification, Dayal (1996) argues that a plural *wh*-question has sum-based answers while a singular *wh*-question doesn’t. For example, in a context where two children came, the plural *wh*-question (60a) 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 (60b) doesn’t have an exhaustive true answer, which violates the EP.

(60) (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 first-order *wh*-quantification. As for higher-order *wh*-quantification, clearly singular *which*-phrases cannot quantify over Boolean conjunctions; otherwise (60b) would have an exhaustive true answer $\lambda w. \text{came}_w(a) \wedge \text{came}_w(b)$ formed out of $a^\uparrow \cap b^\uparrow$ to satisfy the EP.¹⁷

Higher-order *wh*-quantification, which will need to be attended to in modalized questions, exhibits a ‘disjunction–conjunction asymmetry’: the higher-order quantification domain of a singular *which*-phrase excludes Boolean conjunctions but includes Boolean disjunctions (Xiang 2021b). For example in (61), the disjunctive answer to a singular *should*-question may be interpreted with narrow scope relative to *should*. This interpretation argues that the *wh*-domain of *which textbook* contains Boolean disjunctions over atomic books.

(61) 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: as seen in (62), the MA answer to the singular *can*-question can be expressed as an elided FC-disjunction, although not as an elided conjunction.

¹⁷A similar constraint applies if one follows Fox (2018, 2020) and assumes that a *wh*-domain may include higher-order pluralities: singular *which*-phrases cannot quantify over higher-order pluralities, otherwise (60b) would have an exhaustive true answer based on $\{ \{a\}, \{b\} \}$.

- (62) Which textbook can I use for this class?
- Heim & Kratzer* or *Meaning & Grammar*.
 - # *Heim & Kratzer* and *Meaning & Grammar*.

Numeral-modified *which*-phrases as well trigger a uniqueness presupposition, as exemplified in (63). 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 (Xiang 2021b).

- (63) 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 EP

Dayal's EP requires a question to have an exhaustive true answer. In contrast, the max-informativity-based definition of answerhood adopted from Fox predicts that a question has a MS interpretation only if this question can have multiple max-informative true answers (Sect. 4.2.3); in other words, a MS interpretation is only available in cases where Dayal's EP is not trivially satisfied. This is where the dilemma arises: Dayal's EP 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 capture uniqueness.

Despite the status our model accords to MS, abandoning Dayal's EP would be too costly in my view. In addition to explaining the uniqueness effects in questions, this 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 Stechow et al. 2014). Moreover, it can rule out many unwanted question interpretations; without further restrictions, abandoning Dayal's EP 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, 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.

- (64) *Which child came? / Which two children formed a team? / Did you invite Andy, Billy, or Cindy?*
- With Dayal's EP: MA with uniqueness ✓
 - Without Dayal's EP: MS without uniqueness ✗

Case 2: *Wh*-questions with a non-distributive predicate. For *wh*-questions with a non-distributive predicate, the individual answers are logically independent even if the *wh*-phrase is number-neutral or a bare plural. As argued 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, Fox's concept of answerhood predicts a MS interpretation for first-order *wh*-quantification.

- (65) *Which children formed a team?*
- With Dayal's EP: MA without uniqueness ✓

- b. Without Dayal’s EP: Ambiguous between MS and MA ✗

Case 3: *Wh*-questions with an indefinite. *Wh*-questions with an indefinite have choice interpretations and individual interpretations. The choice interpretation in (66a) requests the specification of a boy-movie(s) pair, whereas the individual interpretation in (66b) requests the specification of some movies, but not the names of the boys who watched these movies.

- (66) (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 focuses more on the choice interpretation. Some works (George 2011; Fox 2013; Nicolae 2013; a.o.) treat the choice interpretation as a MS interpretation and claim that indefinites can license MS. However, a number of empirical distinctions argue that choice interpretations and MS interpretations should be treated separately.¹⁸ Here, let’s switch our focus to the individual interpretation, in which the indefinite is interpreted inside the local IP. In principle, (66) could have the LFs (67a,b), which yield interpretations that allow this question to have multiple max-informative true answers. Compared to the LFs assumed for the MS interpretations of *can*-questions, the only difference is that these LFs involve an existential quantifier, as opposed to an existential modal. Once Dayal’s EP is abandoned, it is unclear how we can rule out these LF options.

- (67) a. $[_{CP} \text{ which-movie(s)} \lambda y_e [_{IP} \text{ one-boy } \lambda x_e [_{OC} [_{VP} x \text{ watched } y]]]]$ (cf. (34))

- b. $[_{CP} \text{ which-movie(s)} \lambda \pi_{\langle et, t \rangle} [_{IP} \text{ one-boy } \lambda x_e [_{\pi} \lambda y_e [_{OC} [_{VP} x \text{ watched } y]]]]]]$ (cf. (37))

¹⁸Dayal (2016: Sect. 3.2.1) provides two reasons for keeping choice interpretations separate from MS interpretations. One revolves around morphological differences among *wh*-expressions, and the other is concerned with the availability of MS interpretations in questions with a universal quantifier. Let me add two additional arguments here.

First, choice questions do not have FC-disjunctive answers. In (i), the disjunctive answer only has an ignorance reading. The unavailability of a FC reading argues that the machinery that makes a *can*-question congruent with FC-disjunctive answers is unavailable 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 indefinite in a choice question scopes fairly high at LF (Groenendijk and Stokhof 1984; Chierchia 1993; Szabolcsi 1997; Krifka 2001; Dayal 2016; Xiang 2022; a.o.) and therefore cannot interact with an anti-exhaustification operator.

- (i) a. Which movie did one of the two boys watch? (Choice)
 b. Andy watched *Hulk*, or Billy watched *Ironman*.
 (Doesn’t imply: ‘Andy watched *Hulk*, and Billy watched *Ironman*.’)

Second, the two types of questions behave 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 (83)–(86) and (105)/(107)). In choice questions, however, the local uniqueness presupposition is existential. For example, the question in (i) implies that one of the boys watched exactly one movie, not that each boy watched at most one movie. This contrast argues that uniqueness is derived differently in these two types of questions.

5.1.3. A challenge to Dayal’s EP: local uniqueness

Hirsch and Schwarz (2020) observe that the uniqueness inference in a singular \diamond -question can 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 EP is applied to the answer space as a whole, it cannot account for these local-uniqueness inferences.

- (68) 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: the modal verb *could* is a priority modal in (68a) and an epistemic modal in (68b). Let me note here that local uniqueness is also independent of modal force. For example, the singular *have to*-question in (69) also has a local-uniqueness inference.

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

Dayal’s EP cannot account for the local-uniqueness inference in (69). 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 (70a). Next, assume the scenario in (70b), 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 EP is not violated.

- (70) Which chapter do we have to assign?
 $[_{CP} \text{ which-chapter } \lambda \pi_{\langle et, t \rangle} [_{IP} \text{ have-to } [\pi \lambda x_e [_{VP} \text{ we assign } x]]]]]$
- a. $Q = \{ \Box\phi_{c_1}, \Box\phi_{c_2}, \Box(\phi_{c_1} \vee \phi_{c_2}) \}$ $(\phi_x$ abbreviates ‘we assign x ’)
- 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, referred to as the partition-by-exhaustification (PbE) analysis. Fox first re-writes Dayal’s definition of answerhood in terms of exhaustification, as in (71): 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.

- (71) $\text{ANS}_{\text{Dayal}} := \lambda w \lambda Q : \exists p \in Q [O_Q(p)(w) = 1] . \lambda p \in Q [O_Q(p)(w) = 1]$
 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 O -operator with one that may trigger FC, namely the exhaustivity operator $O^{\text{IE+II}}$ from Bar-Lev and Fox 2020. As defined below, this operator negates the innocently excludable (IExcl-)alternatives and affirms the innocently includable (IIincl-)alternatives.

$$(72) \quad O_C^{\text{IE+II}} = \lambda p \lambda w. \underbrace{\forall q \in C [q \in \text{IExcl}(p, C) \rightarrow q(w) = 0]}_{\text{Innocent Exclusion (IE)}} \wedge \underbrace{\forall r \in C [r \in \text{IIincl}(p, C) \rightarrow r(w) = 1]}_{\text{Innocent Inclusion (II)}}$$

- a. $\text{IExcl}(p, C)$
 $= \bigcap \{A \subseteq C \mid A \text{ is a maximal subset of } C \text{ s.t. } \{\neg q \mid q \in A\} \cup \{p\} \text{ is consistent}\}$
- b. $\text{IIincl}(p, C)$
 $= \bigcap \{B \subseteq C \mid B \text{ is a maximal subset of } C \text{ s.t. } B \cup \{p\} \cup \{\neg q \mid q \in \text{IExcl}(p, C)\} \text{ is consistent}\}$

Let's see how the $O^{\text{IE+II}}$ -operator derives FC and predicts its distribution. In (73), 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 (74), 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 scopes below an existential modal.

$$(73) \quad \text{Let } p = \diamond(\phi_a \vee \phi_b) \text{ and } C = \{\diamond\phi_a, \diamond\phi_b, \diamond(\phi_a \vee \phi_b), \diamond(\phi_a \wedge \phi_b)\}; \text{ then we have:}$$

- a. $\text{IExcl}(p, C) = \{\diamond(\phi_a \wedge \phi_b)\}$ and $\text{IIincl}(p, C) = \{\diamond\phi_a, \diamond\phi_b, \diamond(\phi_a \vee \phi_b)\}$
- b. $O_C^{\text{IE+II}}(p) = \underbrace{\neg\diamond(\phi_a \wedge \phi_b)}_{\text{IE}} \wedge \underbrace{[\diamond\phi_a \wedge \diamond\phi_b \wedge \diamond(\phi_a \vee \phi_b)]}_{\text{II}} = \neg\diamond(\phi_a \wedge \phi_b) \wedge \diamond\phi_a \wedge \diamond\phi_b$

$$(74) \quad \text{Let } p = \phi_a \vee \phi_b \text{ and } C = \{\phi_a, \phi_b, \phi_a \vee \phi_b, \phi_a \wedge \phi_b\}; \text{ then we have:}$$

- a. $\text{IExcl}(p, C) = \{\phi_a \wedge \phi_b\}$ and $\text{IIincl}(p, C) = \{\phi_a \vee \phi_b\}$
- b. $O_C^{\text{IE+II}}(p) = \underbrace{\neg[\phi_a \wedge \phi_b]}_{\text{IE}} \wedge \underbrace{[\phi_a \vee \phi_b]}_{\text{II}}$

Dayal's EP can now be modified to the following:

(75) *Modified EP* (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 $O_Q^{\text{IE+II}}(p)$ is true in w .

Just like Dayal's EP, the modified EP is trivially satisfied if the answer space is closed under conjunction. However, the modified EP can also be trivially satisfied in cases where the answer space contains disjunctions that can be strengthened into FC statements via the application of $O^{\text{IE+II}}$. This property allows for MS. For a concrete example, consider the *can*-question in (76). With three relevant individuals a, b, c , the answer space of the question is as laid out in (76a). Applying $O_Q^{\text{IE+II}}$ point-wise to this answer space returns a set of mutually exclusive propositions as in (76b), referred to as 'the partition induced by Q' '. The modified EP is satisfied in any world that is in the union of this partition, including worlds with multiple possible chair persons.

$$(76) \quad \text{Who can chair the committee alone?} \quad (\phi_x \text{ abbreviates 'x chairs the committee alone'})$$

- a. $Q = \left\{ \begin{array}{l} \diamond\phi_a, \diamond\phi_b, \diamond\phi_c \\ \diamond(\phi_a \vee \phi_b), \diamond(\phi_b \vee \phi_c), \diamond(\phi_a \vee \phi_c) \\ \diamond(\phi_a \vee \phi_b \vee \phi_c) \end{array} \right\}$

$$\begin{aligned}
& \text{b. } \{O_Q^{\text{IE}+\text{II}}(p) \mid p \in Q\} \\
& = \left\{ \begin{array}{l} \diamond\phi_a \wedge \neg\diamond\phi_b \wedge \neg\diamond\phi_c, \diamond\phi_b \wedge \neg\diamond\phi_a \wedge \neg\diamond\phi_c, \diamond\phi_c \wedge \neg\diamond\phi_a \wedge \neg\diamond\phi_b \\ \diamond\phi_a \wedge \diamond\phi_b \wedge \neg\diamond\phi_c, \diamond\phi_b \wedge \diamond\phi_c \wedge \neg\diamond\phi_a, \diamond\phi_a \wedge \diamond\phi_c \wedge \neg\diamond\phi_b \\ \diamond\phi_a \wedge \diamond\phi_b \wedge \diamond\phi_c \end{array} \right\}
\end{aligned}$$

This analysis also avoids over-generating MS interpretations for questions with a non-distributive predicate. For example, in a multiple-team scenario, question (77) (see also (39)) satisfies the modified EP 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 (as assumed in Fox 2018, 2020, discussed in Sect. 4.2.3) or Boolean conjunctions (as assumed in my account).¹⁹

- (77) a. Which children formed a team?
b. Andy and Billy formed a team, and Cindy and Danny formed a team.

What about uniqueness? At first sight, the modified EP should preserve the merits of Dayal’s EP in explaining uniqueness: if a question doesn’t have FC answers, the modification to Dayal’s EP shouldn’t affect the predictions. However, in actuality this problem remains unsolved because the $O^{\text{IE}+\text{II}}$ -operator over-generates FC-disjunctive answers to non-modalized singular *wh*-questions.

Consider the singular *wh*-question in (78). With two boys *a, b* in the discourse domain, the answer space is Q_{FO} in the first-order interpretation and Q_{HO} in the higher-order interpretation. In (78b), applying point-wise exhaustification to Q_{FO} induces a partition that covers only the worlds where exactly one of the boys came, which nicely predicts uniqueness. However, in the higher-order interpretation, applying $O^{\text{IE}+\text{II}}$ to the disjunctive answer $\phi_a \vee \phi_b$ would strengthen it into a FC statement: distinct from the alternative set *C* in (74), 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. $\phi_a \vee \phi_b$. Hence in (78c), the partition induced by Q_{HO} covers also the worlds where both boys came, which therefore predicts no uniqueness effect.

- (78) 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 $\text{IIINCL}(\phi_a \vee \phi_b, Q_{\text{HO}}) = \{\phi_a \vee \phi_b, \phi_a, \phi_b\}$;
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 indefinite. Section 5.1.2 has argued that the answer space of (79) couldn’t be the one given in (79a), since this would allow for a non-exhaustive individual interpretation. However, applying $O_Q^{\text{IE}+\text{II}}$ point-wise to this answer space yields the partition in (79b), which includes worlds where two movies were watched, each by a different boy.

¹⁹The analysis proposed in Fox 2018, 2020 is more complex than what is presented here. To account for negative island effects in higher-order interpretations, Fox argues for a *non-vacuity* principle: 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 non-modalized *wh*-questions do not have higher-order interpretations: in a higher-order interpretation, the answer space of a non-modalized *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 (77), Fox further assumes that the quantification domain of *which children* includes higher-order pluralities such as $\{\{a, b\}, \{c, d\}\}$.

- (79) Which movie(s) did one of the boys watch? (Individual)
- a. $Q = \{\exists x.\phi_{m_1}^x, \exists x.\phi_{m_2}^x, \exists x[\phi_{m_1}^x \vee \phi_{m_2}^x]\}$ (ϕ_m^x abbreviates ‘boy x watched m ’)
 - b. $\{O_Q^{\text{IE}+\text{II}}(p) \mid p \in Q\} = \{\exists x.\phi_{m_1}^x \wedge \neg\exists x.\phi_{m_2}^x, \exists x.\phi_{m_2}^x \wedge \neg\exists x.\phi_{m_1}^x, \exists x.\phi_{m_1}^x \wedge \exists x.\phi_{m_2}^x\}$
 [N.B.: $\text{IExCL}(\exists x[\phi_{m_1}^x \vee \phi_{m_2}^x], Q) = \emptyset$ and $\text{IInCL}(\exists x[\phi_{m_1}^x \vee \phi_{m_2}^x], Q) = Q$;
 therefore: $O_Q^{\text{IE}+\text{II}}(\exists x[\phi_{m_1}^x \vee \phi_{m_2}^x]) = \exists x.\phi_{m_1}^x \wedge \exists x.\phi_{m_2}^x$]

In sum, Fox’s PbE analysis nicely allows for MS; however, since the $O^{\text{IE}+\text{II}}$ -operator over-generates FC, 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; on my account, the partition could be induced by applying both O^{IE} and DOU to each propositional answer. Fox (2018, 2020) doesn’t deal with local-uniqueness effects. For a recent PbE-based analysis of local uniqueness, see Kobayashi and Rouillard 2021.

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

As mentioned, 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 (80) and assume that *which* is interpreted within the question nucleus.

$$(80) \llbracket \textit{which} \rrbracket = \lambda x_e \lambda m_{\langle e, st \rangle} \lambda g_{\langle e, st \rangle} \lambda w : \exists! y [f(y)(w) \wedge g(y)(w)]. f(x)(w) \wedge g(x)(w)$$

Hirsch and Schwarz further argue that global/local uniqueness arises if the *which*-phrase scopes above/below the existential modal, as exemplified in (81). Here each propositional answer carries a global/local uniqueness presupposition (underlined). For the local-uniqueness interpretation (81b), if none of the accessible worlds satisfies uniqueness, the question has no true answer and is deviant.

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

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 below. Here the presupposition of *which* is trivially satisfied when the intersection between the *wh*-domain and the extension of the predicate that the *which*-phrase combines with is closed under sum, such as in *Which students came?*.

$$(82) \llbracket \textit{which} \rrbracket = \lambda x_e \lambda m_{\langle e, st \rangle} \lambda g_{\langle e, st \rangle} \lambda w : \exists y [f(y)(w) \wedge g(y)(w) \wedge \forall z [f(y)(w) \wedge g(y)(w) \rightarrow z \leq y]]. f(x)(w) \wedge g(x)(w)$$

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

However, the presuppositional-*which* account also faces several problems. First, this account only deals with uniqueness effects; it cannot avoid over-generating MS interpretations for *wh*-questions with a non-distributive predicate or with an indefinite.

Second, it 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. Consider (83). 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* is satisfied, and further, that any true answer that specifies a single chapter is a good MS answer here. However, the continuation in (83) is clearly unacceptable: the uniqueness requirement is not satisfied in worlds where we assign "chapter 3".²⁰

(83) (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*, the predicted local-uniqueness inference is read as "We are allowed to assign a single chapter", not "We are *only* allowed to assign a single chapter". However, (84a) 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 *SINgle*, as in (84b), or by using the number-neutral word *what*, as in (84c). (For an explanation of the contrast between (84a) and (84b), see Sect. 6.3.3.)

(84) (As in (83).)

- a. ? I know which chapter we can assign ...,
chapter 1/ chapter 2/ chapter 1 or chapter 2.
- b. I know which SINgle chapter we can assign ...,
chapter 1/ chapter 2/ chapter 1 or chapter 2.
- c. I know what we can assign.

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 verifies a true answer. Dialogue (85) 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 (86) 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 was asking is simply a sub-question of the more general question 'Which journal article or book chapter should we assign?', namely, 'If we choose to assign a journal article, what could it be?'

(85) Instructor: 'We should assign one or two papers.'

TA1: "Got it. Which paper could we assign?"

(Unnatural)

²⁰Oddness arises even before the answer continuation "chapter 3" is given, due to the problem we turn to next.

- (86) Instructor: ‘We should assign a journal article or a book chapter.’
 TA2: “Got it. Which journal article could we assign?” (Natural)

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 (87b), 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.

- (87) Which boy watched which movie?
 a. Single-pair: ‘Which unique boy- x -movie- y pair is such that x watched y ?’
 b. Pair-list: ‘[Each boy watched at most one movie; tell me:] which boy- x -movie- y pairs are such that x watched y ?’

By contrast, point-wise uniqueness in pair-list interpretations can be derived based on Dayal’s EP (Dayal 1996, 2016; Fox 2012; Xiang 2022).

6. Solving the dilemma: Relativized Exhaustivity

Section 5 has presented a dilemma: Dayal’s EP is incompatible with MS, but abandoning it would leave uniqueness effects unexplained and would over-generate MS interpretations for a variety of question types. Further, the presented observations on local uniqueness suggest that Dayal’s EP 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 EP that correctly predicts the distribution of both 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 utterance world. For example, for the modalized question in (88a), local uniqueness says that in every accessible world at most one chapter is assigned. This inference is formally stated in (88b), where w is the utterance world, and M is the modal base of *can/have to* which maps the utterance world to a set of accessible worlds.

- (88) a. Which chapter can we/ do we have to assign?
 b. $\forall w' \in M_w[\text{we assign any chapters in } w' \rightarrow \text{we assign only one chapter in } w']$

Further, in light of Dayal’s insight that uniqueness comes from exhaustivity, I argue that question interpretations do require exhaustivity, but in modalized questions exhaustivity is evaluated relative to the accessible worlds, not the utterance world. More generally, modalized questions are subject to *Relativized Exhaustivity (RelExh)*, described as follows:

- (89) A modalized question ΔQ (where Δ is a modal) is defined only if its non-modalized counterpart Q satisfies Dayal’s EP in every accessible world where Q has a true answer.

How can we capture this idea analytically in model-theoretic compositional semantics? Clearly, this cannot be achieved by changing the interpretation world from the utterance world to a world

accessible to the utterance world: what we want is to evaluate the exhaustivity of the corresponding non-modalized question Q , not that of the modalized question ΔQ itself. However, technically, we cannot extract Q from the meaning of ΔQ . Alternatively, in what follows, this idea is captured by changing the modal base: interpreting Q in an accessible world w' is equivalent to interpreting ΔQ w.r.t. a modal base that introduces only the singleton set of accessible worlds $\{w'\}$.

To make the question denotation sensitive to the modal base, I redefine Dayal's EP as in (90). 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 . Next, the RelExh condition is defined as in (91): Dayal's EP is satisfied relative to every modal base that introduces a singleton set of accessible worlds which verifies a true answer. (**D** and **R** refer to 'Dayal' and 'relativized', respectively. Relevant concepts were introduced in Sect. 4.4.2.)^{21,22}

- (90) *Dayal's EP* (adapted from Dayal 1996)
 Given a modal base M , a question Q is defined in w only if
 $\exists \alpha [\alpha \in \llbracket Q \rrbracket_w^M \wedge \forall \beta [\beta \in \llbracket Q \rrbracket_w^M \rightarrow \llbracket Q \rrbracket^M(\alpha) \subseteq \llbracket Q \rrbracket^M(\beta)]]$ [abbreviated as **DEP**($w, M, \llbracket Q \rrbracket$)]
- (91) *Relativized Exhaustivity*
 Given a modal base M , a question Q is defined in w only if
 $\forall M'_{(s,st)} [\lvert M'_w \rvert = 1 \wedge M'_w \subseteq M_w \wedge \exists \alpha [\alpha \in \llbracket Q \rrbracket_w^M \wedge \alpha \in \llbracket Q \rrbracket_w^{M'}] \rightarrow \mathbf{DEP}(w, M', \llbracket Q \rrbracket)]$
 '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 EP in w .' [abbreviated as **REP**($w, M, \llbracket Q \rrbracket$)]

The next two subsections will present the applications and predictions of RelExh. 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 singular *wh*-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 first-order/higher-order MS denotations:

- (92) Who can chair the committee? (ϕ_x abbreviates 'x chairs the committee')
- $\llbracket Q_{\text{FO}} \rrbracket^M = \lambda x_e : \text{hmn}_{@}(x). [\lambda w. \exists w' \in M_w [w' \in O_C \phi_x]]$
 - $\llbracket Q_{\text{HO}} \rrbracket^M = \lambda \pi_{(et,t)} : \pi \in {}^{\text{H}}\text{hmn}_{@}. [\lambda w. \exists w' \in M_w [w' \in \pi(\lambda x_e. O_C \phi_x)]]$

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 (93).

²¹ RelExh is defined based on short answers because the meaning of a short answer is modal independent. In contrast, the meaning of a sentential answer to a modalized question varies as a function of the modal base.

²² A reviewer points out that the definition of RelExh in (91) requires access to the modal base M , which is difficult to obtain from a compositional perspective. One way to solve this problem is to define modalized sentences as functions from modal bases to propositions, not as open propositions with a free modal base variable. This analysis allows the modal base to be retrieved as an argument of the sentential denotation. It also allows RelExh to be defined based on sentential answers, since sentential denotations are arrived at modal independently.

$$(93) \text{ Let chair-the-committee} = \left[\begin{array}{l} w_1 \rightarrow \{a\} \\ w_2 \rightarrow \{b\} \\ \dots \end{array} \right] \text{ and } \left(\begin{array}{l} M_w = \{w_1, w_2\} \\ M_w^1 = \{w_1\} \\ M_w^2 = \{w_2\} \end{array} \right)$$

Given the modal base M , the first-order MS interpretation (92a) 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. Neither is an exhaustive true answer (ETA), and this violates Dayal's EP. 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 exhaustive true answer in w ; likewise for M^2 . Hence, given M , (92a) satisfies RelExh in w .

(94) For the first-order MS interpretation (92a), we have:

$$\begin{array}{ll} \llbracket \text{Q}_{\text{FO}} \rrbracket_w^M = \{a, b\} & \text{ETA: non-existent} \\ \llbracket \text{Q}_{\text{FO}} \rrbracket_w^{M^1} = \{a\} & \text{ETA: } \diamond O_C \phi_a \\ \llbracket \text{Q}_{\text{FO}} \rrbracket_w^{M^2} = \{b\} & \text{ETA: } \diamond O_C \phi_b \end{array}$$

This analysis also applies to the higher-order MS interpretation (92b). The only difference is that (92b) allows for a disjunctive answer formed based on $a^\uparrow \cup b^\uparrow$. However, this answer doesn't affect exhaustivity, since it remains partial regardless of modal base.

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

(95) *Generalization of RelExh to MS interpretations*

The MS interpretations of $\lceil Wh\text{-}A \text{ can } P? \rceil$ satisfy RelExh if and only if the interpretations of the non-modalized question $\lceil Wh\text{-}A \text{ } P? \rceil$ satisfy Dayal's EP 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 $REP(w, M, \llbracket Q \rrbracket)$ and $MaxI(\alpha, w, M, \llbracket Q \rrbracket)$, respectively. Applying an answerhood operator returns a set of max-informative true answers and triggers a RelExh presupposition.²³

$$(96) \text{ } MaxI(\alpha, w, M, \llbracket Q \rrbracket) = 1 \text{ iff } \alpha \in \llbracket Q \rrbracket_w^M \text{ and } \forall \beta \in \llbracket Q \rrbracket_w^M [\llbracket Q \rrbracket^M(\beta) \not\subseteq \llbracket Q \rrbracket^M(\alpha)].$$

(97) Answerhood operators for complete true answers (modified from (27))

- a. $ANS^S(w)(M)(\llbracket Q \rrbracket^M) = \{\alpha \mid MaxI(\alpha, w, M, \llbracket Q \rrbracket)\}$, defined only if $REP(w, M, \llbracket Q \rrbracket)$
- b. $ANS^P(w)(M)(\llbracket Q \rrbracket^M) = \{\llbracket Q \rrbracket^M(\alpha) \mid MaxI(\alpha, w, M, \llbracket Q \rrbracket)\}$, defined only if $REP(w, M, \llbracket Q \rrbracket)$

6.2.2. Avoiding over-predictions of MS

As mentioned in Sects. 4.2.3 and 5.1.2, in questions with a non-distributive predicate like (98a,b), applying Fox's concept of answerhood without further constraints over-predicts a MS interpretation, because the answer spaces of these questions in first-order *wh*-quantification are not closed under conjunction. The RelExh presupposition solves this MS over-generation problem: in a multiple-team scenario, the first-order interpretations of (98a,b) violate RelExh, just as they violate Dayal's EP.

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

²³In light of Uegaki 2021, to account for the projection of the uniqueness presupposition in embeddings, we can assume that the RelExh presupposition is carried by each max-informative true answer, rather than by the answer set as a whole.

Moreover, RelExh explains why indefinites do not license MS: RelExh relativizes exhaustivity to a smaller modal base, not a smaller discourse domain. The essence of the analysis is that language users naturally shift the origo for interpretation to worlds where the state under discussion actually emerges.²⁴ Evaluating exhaustivity w.r.t. singleton modal bases is just an analytical move to capture the perspectival shift. In contrast, the discourse domain is independent of perspective. Hence, RelExh makes the same prediction as Dayal’s EP for any non-modalized questions, including (66) *Which movie(s) 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 singular *wh*-questions. I will consider three types of questions with distinct modal force: (i) non-modalized questions (Sect. 6.3.1), (ii) *have to*-questions (Sect. 6.3.2), and (iii) *can*-questions, interpreted as either MS (Sect. 6.3.3) or disjunctive MA (Sect. 6.3.4). For these 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 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 interpretations. The following exemplifies the uniqueness inferences of modalized singular *wh*-questions and lists the relevant types of question interpretations.²⁵ As shown, this analysis can explain various uniqueness inferences, except global-uniqueness inferences in singular *can*-questions.

- (99) Singular *have to*-questions
 - a. ‘Which chapter do we have to assign?’
 - i. *There is only one chapter that we have to assign.* (Global uniqueness)
 - ii. *We have to assign at most one chapter.* (Local uniqueness)
 - b. First-order (G); wide-scope higher-order (G); narrow-scope higher-order (L)
- (100) Singular *can*-questions without local exhaustification
 - a. ‘Which chapter can we assign?’
 - i. *There is only one chapter that we can assign.* (Global uniqueness)
 - ii. *We can assign one chapter, but not more.* (Universal local uniqueness)
 - b. First/higher-order MS (L); wide/narrow-scope disjunctive MA (L)
- (101) Singular *can*-questions with local exhaustification
 - a. ‘Which SINGle chapter can we assign?’
 - i. *There is only one chapter that we can assign.* (Global uniqueness)

²⁴Relatedly, given the assumption in situation semantics that propositions can be characterized in terms of situations, we may also evaluate exhaustivity w.r.t. minimal situations. This option allows RelExh to be applied to non-modalized questions like *Who has got a light?*, which share many MS-related properties with *can*-questions (van Rooij 2004). For this question, MS is available if exhaustivity can be evaluated w.r.t. minimal situations that exemplify the proposition *Someone has got a light*.

²⁵For example in (99b), ‘first-order (G)’ means that the first-order interpretation of (99a) yields global (G) uniqueness, and ‘narrow-scope higher-order (L)’ means that the narrow-scope higher-order interpretation of (99a) yields local (L) uniqueness.

- ii. *We can assign exactly one chapter singly, aside from possibly available options of assigning more than one chapter simultaneously.* (‘Existential’ local uniqueness)
- b. First/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 EP in explaining the uniqueness effects in (64). Moreover, compared to Hirsch and Schwarz (2020), it also allows for the point-wise uniqueness effect of (87).

- (64) Which child came? (Singular *wh*-question)
- Which two children formed a team? (Numeral-modified *wh*-question)
- Did you invite Andy, Billy, or Cindy? (As an alternative question)
- (87) Which boy watched which movie? (Pair-list multiple-*wh* question)

6.3.2. Uniqueness effects in *have to*-questions

For *have to*-questions (or any \Box -questions, such as *must/should*-questions)²⁶ with a first-order interpretation, RelExh yields the same prediction as Dayal’s EP. For example in (102), with first-order *wh*-quantification, the question calls for an answer naming an atomic chapter and presupposes that there is only one chapter that we have to assign. Dayal’s EP is violated if there are multiple chapters that are assigned in every accessible world. Since the uniqueness requirement is not satisfied in any accessible worlds, the violation of Dayal’s EP 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.

- (102) Q: Which chapter do we have to assign?
 \rightsquigarrow *There is a unique chapter that we have to assign.* (Global uniqueness)
- A: Chapter 1.

However, when the same singular *have to*-question has a narrow-scope higher-order interpretation, RelExh and Dayal’s EP make different predictions. In (103), the question is exhaustively addressed by a narrow-scope disjunction, read as ‘There is no particular chapter that we have to assign — we just need to choose between chapter 1 and chapter 2.’ (Spector 2007, 2008; Xiang 2021b; see Sect. 4.1.3). In this interpretation, the question implies local uniqueness.

- (103) Q: Which chapter do we have to assign? (= (69))
 \rightsquigarrow *We shouldn’t assign more than one chapter.* (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 (104) (cf. (70b)), 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 EP cannot explain local uniqueness: in the local-uniqueness-violating scenario w , the *have to*-question in (103) has only one true answer $\Box(\phi_{c_1} \vee \phi_{c_2})$, which is also its exhaustive true answer.

$$(104) \text{ assign} = \left[\begin{array}{l} w_1 \rightarrow \{c_1\}, w_2 \rightarrow \{c_2\}, \\ w_3 \rightarrow \{c_1, c_2\}, \dots \end{array} \right], \quad M = \left[\begin{array}{l} w \rightarrow \{w_1, w_2, w_3\} \text{ (with uniq.-violation)} \\ w' \rightarrow \{w_1, w_2\} \text{ (without uniq.-violation)} \end{array} \right]$$

²⁶The discussion here is centered on *have to*-questions simply because the modal force of *have to* is unambiguously universal.

In contrast, RelExh is violated in w : since $\{w_3\}$ verifies the true answer $\Box(\phi_{c_1} \vee \phi_{c_2})$, RelExh is satisfied in w only if the question has an exhaustive true answer in w when interpreted relative to a modal base M' such that $M'_w = \{w_3\}$; however, given any such M' , the question has three true answers in w (i.e., $\Box\phi_{c_1}$, $\Box\phi_{c_2}$, and $\Box(\phi_{c_1} \vee \phi_{c_2})$) but no exhaustive true answer.

In sum, for a singular \Box -question, both RelExh and Dayal's EP can predict a global-uniqueness inference. 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 EP. If RelExh were optional, or if it only came into play in cases where Dayal's EP 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 (95), the MS interpretations of the *can*-question (105b) satisfy RelExh if and only if the non-modalized question (105a) has a unique true answer in every accessible world where we assign any chapter(s). This condition is the desired universal local-uniqueness inference.

- (105) a. Which chapter do we assign?
 \rightsquigarrow *We assign exactly one chapter.* (Uniqueness)
- b. Which chapter can we assign?
 \rightsquigarrow *We can assign exactly one chapter, but not more.* (Universal local uniqueness)

For illustration, let's interpret (105b) relative to the modal base M assumed in (104) above. In the first-order MS interpretation, the question has two true answers in w , i.e., $\Diamond\phi_{c_1}$ and $\Diamond\phi_{c_2}$. Since both answers are verified by $\{w_3\}$, RelExh is satisfied in w only if the question has an exhaustive true answer when interpreted relative to a modal base M' such that $M'_w = \{w_3\}$. However, given any such M' , this question has the same set of true answers in w , i.e., $\{\Diamond\phi_{c_1}, \Diamond\phi_{c_2}\}$, which doesn't contain an exhaustive member. Hence, RelExh is violated in w . In contrast, RelExh is satisfied in w' since w' has no uniqueness-violating accessible world like w_3 . The same idea applies to the higher-order MS interpretation.

In (84b), repeated as (106), we saw that the local-uniqueness inference appears to be existential when the *wh*-complement is modified by *SINgle*: (106) requests the addressee to name one chapter which can be the unique chapter to assign, but it does not rule out the possibility that a set containing more than one chapter could be assigned.

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

Why does this local-uniqueness inference appear to be existential? Due to the modifier *SINgle*, the question nucleus is parsed with local exhaustification: 'Which chapter x is such that we can assign *only* x ?' Given the modal base M described in (104), the set of true MS answers to this question in w is $\{\Diamond\phi_{c_1}, \Diamond\phi_{c_2}\}$. The RelExh presupposition predicts the following condition: the non-modalized exhaustified question *Which chapter x is such that we assign only x ?* has a unique true answer in every accessible world where there is a chapter x such that we only assign x . This condition is fairly weak, because 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.

A remaining issue is that RelExh doesn't generate global-uniqueness readings for *can*-questions. To allow for such readings, we can either assume that Dayal's EP is applied optionally, or else that questions have an optional uniqueness condition, namely, that there can be only one complete true answer.

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

As mentioned in Sect. 5.1.1, singular *can*-questions admit only disjunctive MA, not conjunctive MA. This contrast argues that singular *which*-phrases may quantify over Boolean disjunctions but not Boolean conjunctions. Incorporating this finding into the proposed derivation of disjunctive MA, I propose that a singular *can*-question has a MA interpretation only if an anti-exhaustification operator, i.e., DOU , applies above *can*.

- (107) Which chapter can we assign?
 \rightsquigarrow We can assign exactly one chapter, but not more. (Universal local uniqueness)
- Chapter 1 or chapter 2.
 - # Chapter 1 and chapter 2.

Section 6.3.3 has argued that RelExh can explain the universal local-uniqueness effect of question (107) if this question has a MS interpretation. Now let's see whether the RelExh-based analysis extends to MA interpretations.

The local-uniqueness-violating context is repeated below. In this context, there are options to assign chapter 1 or chapter 2 singly as well as the option to assign the two chapters together.

- (108) $\text{assign} = [w_1 \rightarrow \{c_1\}, w_2 \rightarrow \{c_2\}, w_3 \rightarrow \{c_1, c_2\}, \dots], M_w = \{w_1, w_2, w_3\}$

The answer spaces derived by disjunctive MA interpretations are as in Figures 6a,b, which differ only w.r.t. 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.

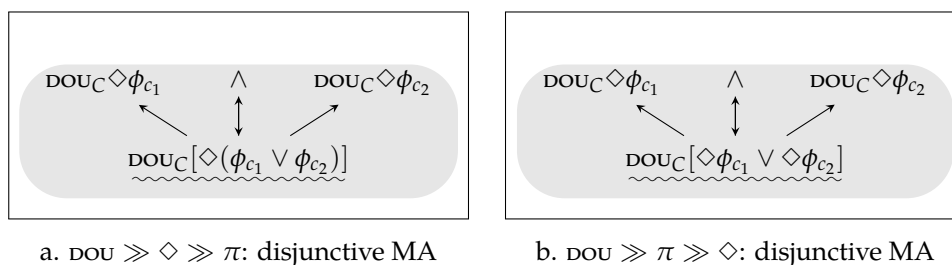


Figure 6: Answer space of (107) in a disjunctive MA interpretation

Now consider the predictions of RelExh. The true answer $\text{DOU}_C \diamond \phi_{c_2}$ (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 a modal base $M'_w = \{w_3\}$. As we saw above, 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' 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' 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 (107) has a disjunctive MA interpretation.

How else then does my account avoid under-generating uniqueness? I'd like to suggest that, even 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 presupposition condition of DOU independently needed to account for the modal obviation effect in the licensing of universal FC. As argued in Sect. 4.4.2, when DOU applies to a \diamond -disjunction, no matter whether the disjunction scopes below or above the existential modal, the RelExcl presupposition of DOU yields a definedness condition as follows:

- (109) DOU_C [John or Mary can teach Calculus]
- 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 (109a), if the preajacent 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 (107), 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 MS interpretations (see (105b)).

This analysis also applies to (110) below, repeated from (106). 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 ?' As in the case of (109b), 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 (110), 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 (106)).

- (110) Q: Which *SINGLE* chapter can we assign?
 \rightsquigarrow *We can assign exactly one chapter 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 DOU -operator. Once the disjunctive answers that violate RelExcl are removed, the RelExh presupposition can account for the universal local-uniqueness effects.

6.4. Section summary

To predict the distribution of MS interpretations and uniqueness effects in questions, this section has assumed a 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 EP 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 EP in deriving uniqueness effects, and avoids over-generating MS interpretations for non-*can*-questions. Moreover, RelExh can account for the local-uniqueness effects of modalized questions. The predictions of Dayal’s EP and the proposed RelExh are summarized as follows. As shown, the only type of uniqueness that RelExh doesn’t derive is global uniqueness in \diamond -questions, which needs to be explained by additional constraints.

Modal type	Reading type	Dayal’s EP	RelExh
No modal	\pm uniq.	✓	✓
\square -modal	– uniq.	✓	✓
	global uniq.	✓	✓
	local uniq.	✗	✓
\diamond -modal	– uniq. MS	✗	✓
	– uniq. MA	✓	✓
	global uniq.	✓	✗
	local uniq. MS	✗	✓
	local uniq. MA	✗	✓

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

7. Conclusions and open issues

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.

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 scopes above the modal *can*, or (ii) an anti-exhaustification operator 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. 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.

This paper is focused on the effect of modal force in licensing MS; I have largely left open the effect of modal flavor and attributed the lack of a MS-licensing effect of epistemic and deontic modals to pragmatics. Relatedly, while I have argued that MS is grammatically licensed in any *can*-question, it can be blocked by external/nucleus-dependent factors, such as contextual factors (Moyer and Syrett 2019) and embedding predicates. We need further empirical and theoretical investigation to fully explain the distribution of MS.

Acknowledgements This paper supersedes chapters two and three of my dissertation (Xiang 2016) and the proceeding papers of Amsterdam Colloquium 20 and NELS 46. The analysis, especially the part on resolving the dilemma between uniqueness and mention-some, has been substantially reworked. I thank Gennaro Chierchia, Danny Fox, and C.-T. James Huang for guiding me during the early development of this work, and two exceptional reviewers of *Natural Language Semantics* and the handling editor Clemens Mayr for extremely helpful reviews. I also thank the audiences at the New York Language of Philosophy Workshop, TaLK 2022 at Keio University, and UConn Logic Group for feedback on the recent versions of this paper. I am specially grateful to Christine Bartels for her incredible editorial comments, which have greatly reshaped the presentation of this paper.

References

- Abrusán, Márta. 2014. *Weak Island Semantics*. Oxford: Oxford University Press.
- Bar-Lev, Moshe E, and Danny Fox. 2020. Free choice, simplification, and innocent inclusion. *Natural Language Semantics* 28:175–223.
- 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. 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. Talk at SuB 18, University of the Basque Country. 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 *Semantics: An international handbook of natural language meaning*, ed., C. Maienborn et al., 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., A. Fălăuș, 88–122. London: Palgrave Macmillan.

- Dayal, Veneeta. 2016. *Questions*. Oxford: Oxford University Press.
- Dotlačil, Jakub, and Floris Roelofsen. 2019. Dynamic inquisitive semantics: Anaphora and questions. In *Proceedings of SuB 23*, ed., M.T. Espinal et al., 365–382.
- 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*, ed., L. Crnic and U. Sauerland, 165–174.
- Fox, Danny. 2007. Free choice disjunction and the theory of scalar implicatures. In *Presupposition and implicature in compositional semantics*, ed., U. Sauerland and P. Stateva, 71–120. New York: Palgrave Macmillan.
- Fox, Danny. 2012. Multiple *wh*-questions: Uniqueness, pair-list and second order questions. MIT seminar. <http://lingphil.mit.edu/papers/fox/uniqueness-pair-list.pdf>.
- Fox, Danny. 2013. Mention-some interpretations. MIT seminar. <http://lingphil.mit.edu/papers/fox/class1-3.pdf>.
- Fox, Danny. 2018. Partition by exhaustification: Comments on Dayal 1996. In *Proceedings of SuB 22*, ed., U. Sauerland and S. Solt, 403–434. Berlin: ZAS.
- 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, UCLA.
- Ginzburg, Jonathan. 1995. Resolving questions, i. *Linguistics and Philosophy* 18:459–527.
- Groenendijk, Jeroen, and Martin Stokhof. 1984. On the semantics of questions and the pragmatics of answers. *Varieties of formal semantics* 3:143–170.
- Heim, Irene. 1994. Interrogative semantics and Karttunen's semantics for *know*. In *Proceedings of IATL 1*, ed., R. Buchalla and A. Mittwoch, 128–144.
- Hirsch, Aron, and Bernhard Schwarz. 2020. Singular *which*, mention-some, and variable scope uniqueness. In *Proceedings of SALT 29*, ed., K. Blake et al., 748–767. Washington, D.C.: LSA.
- 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: de Gruyter.
- 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., Y. Otsu, 1–25. Tokyo: Hituzi Syobo.
- Krifka, Manfred. 2001. 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 SALT 30*, ed., J. Rhyne et al., 714–733. Washington, D.C.: LSA.
- Lahiri, Utpal. 2002. *Questions and answers in embedded contexts*. Oxford: Oxford University Press.
- Li, Haoze. 2021. Mandarin *wh*-conditionals: A dynamic question approach. *Natural Language Semantics* 29:401–451.
- Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, use, and interpretation of language*, ed., C. Schwarze et al., 302–323. Berlin: de Gruyter.

- Liu, Mingming. 2016. Varieties of alternatives: Mandarin focus particles. *Linguistics and Philosophy* 40:61–95.
- Menéndez-Benito, Paula. 2010. On universal free choice items. *Natural Language Semantics* 18:33–64.
- Moyer, Morgan, and Kristen Syrett. (Non-) exhaustivity in embedded questions: Contextual, lexical and structural factors. In *Proceedings of SuB 23*, ed., M.T. Espinal et al., vol 2, 207–224.
- Nicolae, Andreea C. 2013. *Any* questions? Polarity as a window into the structure of questions. Ph.D. Dissertation, Harvard University.
- 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.
- Rullmann, Hotze, and Sigrid Beck. 1998. Presupposition projection and the interpretation of *which*-questions. In *Proceedings of SALT 8*, ed., D. Strolovitch and A. Lawson, 215–232.
- Sharvy, Richard. 1980. A more general theory of definite descriptions. *The Philosophical Review* 89:607–624.
- Spector, Benjamin. 2007. Modalized questions and exhaustivity. In *Proceedings of SALT 17*, ed., M. Gibson and T. Friedman, 282–299. Ithaca, NY: CLC Publications.
- 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. 1997. Quantifiers in pair-list readings. In *Ways of scope taking*, ed., A. Szabolcsi, 311–342. Dordrecht: Kluwer.
- Theiler, Nadine, Floris Roelofsen, and Maria Aloni. 2018. A uniform semantics for declarative and interrogative complements. *Journal of Semantics* 35:409–466.
- Uegaki, Wataru. 2021. The existential/uniqueness presupposition of *wh*-complements projects from the answers. *Linguistics and Philosophy* 44:911–951.
- Xiang, Yimei. 2016. Interpreting questions with non-exhaustive answers. Ph.D. Dissertation. Harvard University.
- 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. 2022. Quantifying into *wh*-dependencies: Multiple-*wh* questions and questions with a quantifier. Manuscript. Rutgers University. <https://ling.auf.net/lingbuzz/005097>.
- Xiang, Yimei, and Alexandre Cremers. 2017. Mentions-some readings of plural-marked questions: Experimental evidence. In *Proceedings of NELS 47*, ed., A. Lamont and K.A. Tetzloff, vol 3, 261–274.
- Zhao, Zhuoye. 2019. Varieties of distributivity: From Mandarin *dou* to plurality, free choice and scalarity. Master thesis. ILLC, University van Amsterdam.